

**LOOKING AT TEACHER LEARNING FROM THE UNIVERSITY:
LEVERS FOR IMPROVEMENT**

by

Hilary Mead

An education leadership portfolio submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Education in Educational Leadership

Spring 2018

© 2018 Hilary Mead
All Rights Reserved

**LOOKING AT TEACHER LEARNING FROM THE UNIVERSITY:
LEVERS FOR IMPROVEMENT**

by

Hilary Mead

Approved: _____
Chrystalla Mouza, Ed.D.
Interim Director of the School of Education

Approved: _____
Carol Vukelich, Ph.D.
Dean of the College of Education and Human Development

Approved: _____
Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this education leadership portfolio and that in my opinion it meets the academic and professional standard required by the University as an education leadership portfolio for the degree of Doctor of Education.

Signed:

Elizabeth Farley-Ripple, Ph.D.
Professor in charge of education leadership portfolio

I certify that I have read this education leadership portfolio and that in my opinion it meets the academic and professional standard required by the University as an education leadership portfolio for the degree of Doctor of Education.

Signed:

Joan Buttram, Ph.D.
Member of education leadership portfolio committee

I certify that I have read this education leadership portfolio and that in my opinion it meets the academic and professional standard required by the University as an education leadership portfolio for the degree of Doctor of Education.

Signed:

Lori Duerr, Ed.D.
Member of education leadership portfolio committee

I certify that I have read this education leadership portfolio and that in my opinion it meets the academic and professional standard required by the University as an education leadership portfolio for the degree of Doctor of Education.

Signed:

Chrystalla Mouza, Ed.D.
Member of education leadership portfolio committee

ACKNOWLEDGMENTS

This Educational Leadership Portfolio was a team endeavor. I am grateful to everyone who participated in the studies and evaluations conducted as part of this ELP. Thank you for allowing me to observe, inquire about, and learn from your experiences with professional learning. This ELP is informed by my work at DASL and DERDC and I am grateful to colleagues in both centers, past and current. Faculty in the School of Education shaped my thinking and built my skills, both in and out of the classroom. I was very fortunate in my committee, all of whom supported and challenged me in unique ways. As my supervisor during much of my degree, Dr. Joan Buttram gave me many opportunities to apply what I was learning about program evaluation and professional learning. Dr. Elizabeth Farley-Ripple continued to believe in me and this ELP, even when I was not sure. I am grateful for her high standards, personal encouragement, and integrity. I thank family and friends who gave me time and space to do this work. As well as being a tremendously supportive spouse, Terence Gilheany is a model of an engaged, curious, constantly learning teacher. My daughters, Hannah and Margaret, fill my life with joy and meaning. I thank them for their patience with “Mommy’s homework” and look forward to all the adventures we can have now that it is done. My mother, sister, and in-laws rounded out the family cheering squad. Thank you all. Finally, my father, Kirtland Chase Mead, died during the last stages of this project, and I dedicate this work to him. I miss you, Dad.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	xiii
ABSTRACT	xiv

Chapter

1	INTRODUCTION.....	1
2	PROBLEM ADDRESSED.....	4
	University Context.....	4
	Delaware Public Education Context.....	6
	Professional Learning Context	7
	Problem Statement.....	9
	Research on Professional Learning	10
3	IMPROVEMENT STRATEGIES.....	16
	Theory of Change	16
	Evaluation of Professional Learning	20
	Five Levers for Improvement.....	25
4	IMPROVEMENT STRATEGIES RESULTS	34
	Conduct research and evaluation related to professional learning.	34
	Synthesize research related to professional learning; develop related recommendations.....	38
	Innovate with evaluation methods.....	39
	Educate others about research-based professional learning design, implementation, and evaluation.....	40
	Increase understanding of what educators value in professional learning efforts and evaluations.....	42
5	REFLECTION ON IMPROVEMENT STRATEGIES RESULTS	45
6	REFLECTION ON LEADERSHIP DEVELOPMENT.....	51
	Scholar	51
	Problem-Solver.....	54
	Partner.....	55

REFERENCES	58
------------------	----

Appendix

A	ARTIFACT A: EVALUATION OF THE 2015 ACTIVITIES OF THE DELAWARE TITLE II GRANT	63
	Artifact.....	64
	Appendix 1: Standards	127
	Appendix 2: Participant demographics	130
	Appendix 3: Sample participant survey	132
	Appendix 4: Participant interview protocol	134
	Appendix 5: Administrator interview protocol	135
	Appendix 6: Writing module checklist.....	136
	Appendix 7: Lesson plan rubric	137
B	ARTIFACT B: STUDY OF THE TEACHER LEADERSHIP INITIATIVE	139
	Artifact.....	139
	Appendix 1: TLI interview protocols.....	187
C	ARTIFACT C: ANNOTATED BIBLIOGRAPHY	191
D	ARTIFACT D: PRACTICE BRIEF, IMPROVING PROFESSIONAL LEARNING IN DELAWARE	227
E	ARTIFACT E: AFTER THE PD: THE ROLE OF SCHOOL LEADERS IN IMPLEMENTATION (NARRATED POWERPOINT)	239
F	ARTIFACT F: DELAWARE DEPARTMENT OF EDUCATION SPECIFIC AND INNOVATIVE IMPROVEMENT GRANT METAEVALUATION.....	259
G	ARTIFACT G: REFLECTIONS FROM TEACHING EDUC 774, SPRING 2016	277
	Artifact.....	277
	Appendix 1: Summary of student evaluation data	289
H	ARTIFACT H: UNDERSTANDING BY DESIGN CURRICULUM UNIT	291
	Artifact.....	291
	Appendix 1: Essential competencies for program evaluators	311
	Appendix 2: Guskey (2000) evaluation levels	313

	Appendix 3: Understanding by Design template.....	315
	Appendix 4: Rubrics for performance tasks.....	323
I	ARTIFACT I: ONLINE PROFESSIONAL DEVELOPMENT TOOLKIT ..	325
J	ARTIFACT J: ANALYSIS OF STUDENT PROFESSIONAL DEVELOPMENT PLANS FROM EDUC 774.....	326
K	ARTIFACT K: EXPLORATION OF EVALUATION USE WITH FORMER CLIENTS	349
	Artifact.....	349
	Interview protocols.....	383
L	PROPOSAL NARRATIVE.....	385
M	INSTITUTIONAL REVIEW BOARD DOCUMENTATION	423

LIST OF TABLES

Table 1	Improvement levers and artifacts	3
Table A1	Overview of Title II program activities 2014-15	68
Table A2	Title II institute enrollments	71
Table A3	District or charter affiliations of Title II participants	72
Table A4	Participant ratings of literacy coaching design.....	73
Table A5	Participants' favorite aspects of literacy coaching.....	74
Table A6	Participants' suggestions for improvement for literacy coaching.....	74
Table A7	Participants' ratings of literacy coaching learning outcomes	75
Table A8	Participants' anticipated challenges for implementing literacy coaching	76
Table A9	Participants' expectations for student response to literacy strategies	77
Table A10	Student learning targets and performance for grade K-2 reading	79
Table A11	Participant ratings of writing module development design.....	81
Table A12	Participants' favorite aspects of writing module development	82
Table A13	Participants' suggestions for improvement for writing module development	82
Table A14	Participants' ratings of writing module development learning outcomes	83
Table A15	Participants' expectations about student response to the writing modules.....	83
Table A16	Participants' anticipated challenges for implementing writing modules .	84
Table A17	Teachers' perception of content knowledge before and after professional development in writing	85
Table A18	Teachers' perception of pedagogy before and after professional development in writing.....	86

Table A19	Distribution of criterion ratings for informative writing modules.....	86
Table A20	Pre and post instruction rubric scores for 3 rd grade students.....	88
Table A21	Pre and post instruction rubric scores for 5 th grade students.....	89
Table A22	Pre and post instruction rubric scores for 9 th grade students.....	89
Table A23	Participants' ratings of summer writing institute design	90
Table A24	Participants' favorite parts of the summer writing institute.....	91
Table A25	Participants' suggestions for changes to the summer writing institute	92
Table A26	Participants' ratings of summer writing institute learning outcomes	92
Table A27	Participants' plans to implement writing strategies into their curricula ..	93
Table A28	Participants' expectations about student reactions to the institute's writing strategies	94
Table A29	Participants' expected challenges to implementation of writing strategies	94
Table A30	Distribution of rubric ratings for participants' draft non-fiction writing lesson plans.....	95
Table A31	Social studies webinar dates, topics, and number attending and registered	96
Table A32	Dissemination of webinar materials	97
Table A33	Participants' ratings of the social studies webinar series design.....	98
Table A34	Participants' favorite parts of the social studies webinar series.....	99
Table A35	Participants' suggestions for changes to social studies webinar series....	99
Table A36	Participants' ratings of online format of social studies webinar series..	100
Table A37	Participants' responses to advantages and disadvantages of online format for social studies webinar series	100
Table A38	Participants' ratings of learning gains in social studies webinar series .	101

Table A39	Participants' ratings of social studies webinar series learning outcomes	102
Table A40	Number and percentage of Responsive Classroom surveys received by school or institute	104
Table A41	Participants' ratings of Responsive Classroom Workshop design.....	105
Table A42	Participants' favorite parts of the Responsive Classroom Workshop....	106
Table A43	Participants' suggestions for changes to the Responsive Classroom Workshop	107
Table A44	Participants' ratings of Responsive Classroom Workshop learning outcomes.....	108
Table A45	Participants' expectations for student response to Responsive Classroom strategies.....	109
Table A46	Participants' expected challenges to implementing Responsive Classroom strategies.....	110
Table A47	School disciplinary action data for 2013-14 and 2014-15.....	112
Table A48	Participants' ratings of Educator Design Workshop design	114
Table A49	Participants' favorite parts of Educator Design Workshop	115
Table A50	Participants' suggested changes to Educator Design Workshop	116
Table A51	Participants' ratings of Educator Design Workshop learning outcomes	116
Table A52	Influence of prior experience on participation in summer design workshop	117
Table A53	Participants' plans to integrate engineering design challenge into their classes or curriculum	118
Table A54	Participants' expectations of student response to Educator Design Workshop strategies	119
Table A55	Participants' anticipated challenges to applying Educator Design Workshop strategies	119
Table A56	Initial members of DARE Cohort 2 measured against admissions criteria.....	121

Table A57	DARE Cohort 1 candidates' ratings of workshops	122
Table A58	DARE Cohort 1 most useful aspects of workshops	122
Table A59	DARE Cohort 1 recommendations for improvement for workshops....	123
Table B1	Study participation and response rate	148
Table F1	SIIP project rubric scores	265
Table F2	Rank order of SIIP projects based on process, outcome, and grand total rubric scores	266
Table F3	Distribution of individual rubric scores.....	266
Table F4	Distribution of scores by rubric	267
Table F5	Distribution of rubric scores by project goal	269
Table J1	Features of professional development	327
Table J2	Alignment with content focus feature	329
Table J3	Alignment with data feature	330
Table J4	Alignment with coherence feature.....	332
Table J5	Alignment with collective participation feature	333
Table J6	Types of collective participation	333
Table J7	Alignment with active learning designs feature	334
Table J8	Types of learning designs	335
Table J9	Alignment with duration feature	336
Table J10	Alignment with follow up feature	337
Table J11	Alignment with leadership feature	338
Table J12	Alignment with differentiation feature	339
Table J13	Alignment with technology feature	341
Table J14	Evaluation methods by Guskey level evaluated	344

Table K1	Participants and projects in the study	357
Table L1	Proposed artifacts	402
Table L2	Table of artifacts.....	415

LIST OF FIGURES

Figure 1	Learning Forward and Delaware Standards for Professional Learning	9
Figure 2	TELL Delaware Data	14
Figure 3	ELP Theory of Change	18
Figure 4	Revised Theory of Change	49
Figure B1	TLI snowflake scaling model	160
Figure C1	Guskey (2002) model	194
Figure C2	Desimone (2009) model	198
Figure C3	Clarke & Hollingsworth (2002) model.....	201
Figure C4	Evans (2014) model.....	207
Figure C5	Coldwell & Simkins (2011) model.....	209
Figure C6	King (2014) model	211
Figure C7	Kennedy (2014) model	217
Figure C8	Timperley (2011) model.....	220
Figure J1	Content focus of PDPs.....	329
Figure J2	PDPs evaluating each Guskey leve	343
Figure L1	Learning Forward and Delaware Standards for Professional Learning	392
Figure L2	TELL Delaware data	397

ABSTRACT

This Educational Leadership Portfolio (ELP) investigates how university staff can contribute to high-quality professional learning in K-12 public schools. The current shortcomings of most professional learning are well documented in research and policy, and in the lived experiences of teachers and administrators. There is an urgent need to redesign professional learning and implement it more effectively, thus supporting improvements in teaching and learning.

I argue that research and evaluation professionals at the University of Delaware (UD) can use five levers to effect positive change in professional learning. They can directly conduct research studies or evaluations of programs and innovate with new methods of doing so. They can synthesize and disseminate research and educate others about research-based professional learning. Finally, to make their research and evaluations useful to K-12 public educators, they can listen to what those users value and need. Together, these activities can yield more informed professional learning leaders and better research or evaluation of professional learning. These outcomes then contribute to stronger professional learning designs, implementation and leadership.

This ELP documents my use of these five improvement strategies as a UD professional, primarily through program evaluations I conducted at the Delaware Education Research & Development Center. There are 11 artifacts illustrating the strategies. These include two evaluations and one exploratory study of professional learning programs; a presentation, a practice brief and a website; an annotated bibliography about teacher learning; three artifacts from formal and informal educational efforts; and a follow-up study with former clients about evaluation use.

After reviewing the results of each improvement lever, this ELP offers the following reflections. I successfully informed professional learning leaders and discovered that two of the five levers were the most effective (i.e., conduct research and evaluation about professional learning; educate others about professional learning and program evaluation). I grew as a researcher and evaluator of professional learning, especially qualitatively, and learned more about making my work useful to practitioners. I also found that evaluation use, like the transfer of professional learning to instructional practice, is highly situated. Organizational culture, resources, and leadership can facilitate or constrain it.

Chapter 1

INTRODUCTION

This Educational Leadership Portfolio (ELP) addresses and unpacks a multi-dimensional question: how can universities best facilitate professional learning for educators in the K-12 public system? The current shortcomings of most professional learning are well documented in research and policy, and in the lived experiences of teachers and administrators. There is an urgent need to redesign professional learning and implement it more effectively, thus supporting improvements in teaching and learning. In this portfolio, I argue that university-based professionals like me have unique opportunities and resources to address this need. Specifically, I identify five levers to effect change in K-12 professional learning. Universities can:

1. Conduct research and evaluation related to professional learning.
2. Synthesize research related to professional learning; develop related recommendations.
3. Innovate with evaluation methods.
4. Educate others about research-based professional learning design, implementation, and evaluation.
5. Increase understanding of what educators value in professional learning efforts and evaluations.

This portfolio investigates the intersections of professional learning, program evaluation, and educational leadership. My integrative approach reflects my broad interests and experiences. I have always been fascinated by the art and technique of teaching, and the challenge of improving it. Since 2006, I have worked at UD in a variety of roles: as a graduate research assistant (2006-08) and associate (2008-11) at

the Delaware Academy for School Leadership (DASL) and as an educational researcher and program evaluator at the Delaware Education Research and Development Center (DERDC) (2012-15). For the past two years, I have conducted part-time projects for both centers, as well served as a preceptor for classes in the School of Education (SoE) and a member of other SoE research teams. My perspective is one of a university professional staff member whose main responsibilities are not teaching but rather conducting research, evaluation, and school improvement projects. I use these tools to strengthen professional learning and engage educational leaders in this process.

The rest of this essay is organized as follows. Chapter 2 establishes the need to improve professional learning, using both Delaware and national data. It also builds state and national context and introduces research about effective professional learning. Chapter 3 digs further into research and program evaluation as a strategy for improving professional learning design and implementation. It then segues to my improvement strategies, the five levers identified above, and my overall theory of change. Chapter 4 examines the effectiveness of these strategies, looking at how and how well I used each lever. Chapter 5 reflects more broadly on the results of my efforts and lessons learned from this project and Chapter 6 discusses my leadership development and career trajectory.

Eleven artifacts are included in this portfolio and demonstrate my efforts to use the five improvement levers discussed above. They are listed in Table 1, overleaf.

Table 1 Improvement levers and artifacts

Improvement Lever	Artifacts
1. Conduct research and evaluation related to professional learning	A. Evaluation of the 2015 Activities of the Delaware Title II Grant B. Study of the Teacher Leadership Initiative
2. Synthesize research related to professional learning; develop related recommendations	C. Annotated bibliography about models of teacher learning D. Practice brief, Improving Professional Learning in Delaware E. Narrated PowerPoint, After the PD: The role of school leaders in implementation
3. Innovate with evaluation methods.	F. Delaware Department of Education Specific and Innovative Practices Grant Meta-evaluation
4. Educate others about research-based professional learning design, implementation, and evaluation	G. Reflections from teaching EDUC 774, Designing Professional Development H. Understanding by Design curriculum unit to teach graduate research assistants about professional learning. I. Online professional development toolkit
5. Increase understanding of what educators value in professional learning and evaluation	J. Analysis of student professional development plans from EDUC 774 K. Exploration of evaluation use with former clients

Overall, my portfolio explores how university professionals can add value to the professional learning that occurs within K-12 schools. How can we translate insights from research to inform stronger programs or practices, helping to achieve better teaching and learning outcomes? How can we ensure that research efforts are grounded in the real world of schools? Conversely, how can a busy district, school or teacher leader integrate a more research-based perspective into his or her responsibilities to plan, deliver, and evaluate professional learning? My portfolio documents how I have addressed such questions.

Chapter 2

PROBLEM ADDRESSED

This chapter establishes the need to improve professional learning in Delaware public schools and the opportunity that the University of Delaware (UD) has to do so. This chapter first broadly describes the university and Delaware public education landscapes, then situates professional learning for educators within those landscapes. It reviews state and national policies and standards related to professional learning and examines how well current practice lives up to them. It also briefly summarizes the research consensus about effective professional learning.

University Context

UD engages with the broader community, including K-12 education, in many ways through its seven colleges. In 2015, the university was awarded the Carnegie Community Engagement Classification in recognition of these efforts. In its Carnegie application, UD demonstrated strong investments in community engagement. For instance, a campus-wide survey showed that faculty committed 26% of their work time to this purpose; for professional staff it was 24% (University of Delaware, 2014). The university also laid out a rationale for community engagement:

UD recognizes that community engagement is critical not only to our public service mission but also to our educational and research missions: enriching student learning, improving the effectiveness of our teaching, allowing partnerships to guide research, and enhancing the impact of scholarship (University of Delaware, 2014, p. 2).

Using this broad framework, UD's partnerships with K-12 public education are seen as mutually beneficial. The ultimate aim of improving student learning outcomes

in Delaware clearly supports UD’s “public service...educational and research missions.” More proximally, by getting involved with K-12 professional learning through strategies such as those discussed in this portfolio, university personnel not only serve the public but build relationships and gain insight into educational contexts which can yield research partnerships, opportunities to develop new interventions or research methods, and audiences for dissemination and application of research. The university and school, district, and/or state capacities are reciprocally strengthened.

There is increasing interest and momentum for such partnerships in Delaware. In 2015, UD unveiled a new strategic plan, which included community engagement as a key initiative and made the following priority recommendation (emphases mine): “Engage more Delaware preK-12 teachers and their classrooms in UD research/ outreach programs, the traditional and performing arts, **teacher professional development**, and innovative curricula” (University of Delaware, 2015). To strengthen and organize existing partnerships between higher and K-12 education as well as establish new ones, in 2016 UD launched the Partnership for Public Education (PPE). While PPE is a relatively new coordinating structure, UD also has numerous research and public service centers that serve as connection points between the university and the broader educational community. Both of the centers for which I have worked, DASL and DERDC, serve this function within the College of Education and Human Development (CEHD).¹ According to its mission statement, CEHD

develops solutions to the problems that confront our schools and the challenges encountered by our children, youth, and families. Although

¹ As of July 1, 2017, DERDC merged with the Center for Research in Education and Social Policy (CRESP). Since work for this portfolio was done for DERDC, I continue to use that name.

our primary mission is to conduct research and train UD students to become highly qualified professionals and leaders in their fields, we also partner with organizations and agencies to ensure that Delaware children, teachers and families receive the best possible education and vital social services (2013).

In 2013, the College estimated that its faculty and staff provided professional development (PD) or training to 10,000 early childcare providers, 2,000 educators and 500 school leaders in Delaware and nearby states. In summary, my ELP's focus on professional learning collaborations between higher education and K-12 is consistent with the organizational structures and strategic direction of both CEHD and UD.

Delaware Public Education Context

I now examine public education in Delaware and demonstrate why the needs for professional learning are so urgent. Delaware has 226 public schools, together enrolling 137,217 students and employing 9,287 teachers and 877 administrators (Delaware Department of Education, 2017a). There is substantial teacher mobility, increasing the need for ongoing development. In the 2016 hiring season, 1,335 teachers were hired meaning more than 14% of Delaware's teachers were new to their schools (Robertson-Kraft, Hoe, Sangenito & Williams, 2017). Statewide, 27.5% of teachers have five or fewer years' experience (Strategic Data Project, 2015).

Delaware education is characterized by standards, assessments, and educator accountability all of which have implications for professional learning. Since 2010, the state has rolled out more rigorous Common Core State Standards (CCSS) for math and English Language Arts (ELA) along with new Smarter Balanced assessments. Delaware also implemented new science standards and an aligned assessment is in development. These changes required massive effort from the entire educational system but no one was more affected than teachers, who had to develop new

instructional practices and often adjust to new curricula as well. Teachers are held accountable for their students' mastery of the standards through the state's appraisal system (DPAS II), one fifth of which is measured by student achievement. Many of these changes are legacies of the state's Race to the Top (RTTT) grant, which invested heavily in "human capital" initiatives related to teacher and leader effectiveness.

When RTTT ended in 2015, performance had improved for some schools and students but gaps remained. Statewide, only 54% of students tested proficient in ELA and 45% in math in 2017, and performance in subgroups (e.g., low-income students, students with disabilities, or English language learners) was lower (Delaware Department of Education, 2017b). Meanwhile, the number of students in these subgroups was growing. Delaware public students are increasingly diverse in culture and language. More and more live in poverty. Recognizing these demographic changes and persistent inequities, the state's Every Student Succeeds Act (ESSA) plan, approved in August 2017, focuses on closing student achievement and educator equity gaps. Doing so requires enhanced educator practices.

Professional Learning Context

In response to these pressing demands, those designing educational programs almost inevitably rely on activities to build educators' knowledge and skills. As Desimone (2009) puts it, "education reform is often synonymous with teachers' professional development" (p. 181). In 2016, the Delaware Department of Education (DDoE) launched a grant competition, *Reimagining Professional Learning*, which acknowledged that "in the past three years, both state and local level approaches to professional learning have evolved in many meaningful ways, and yet there is still so much to do to realize the commitment to reimagined, top-notch professional learning

for every Delaware educator as the norm” (Delaware Department of Education, 2016a). DDoE also put forth this definition:

Professional learning is defined as a comprehensive, sustained, and intensive approach to improving teachers’ and principals’ effectiveness in raising student achievement. There is a shift from the concept of professional development to professional learning to connote the importance of continuous improvement.

Delaware’s definition aligns with the federal criteria in ESSA, which states that professional learning must be “sustained, intensive, collaborative, job-embedded, data-driven and classroom-focused” (Learning Forward, 2015) as well as evidence-based and personalized. Thus professional learning is not a discrete event but a complex process of adult learning, with the purpose of promoting student achievement. Echoing many of these ideas are the Standards for Professional Learning, developed by Learning Forward and adopted in Delaware in 2012.

Delaware's Standards for Professional Learning:	Professional learning that increases educator effectiveness and results for all students...
Standard: Learning Communities	Occurs within learning communities committed to continuous improvement, collective responsibility, and goal alignment.
Standard: Leadership	Requires skillful leaders who develop capacity, advocate, and create support systems for professional learning.
Standard: Resources	Requires prioritizing, monitoring, and coordinating resources for educator learning.
Standard: Data	Uses a variety of data sources and types of student, educator, and system data to plan, assess, and evaluate professional learning.
Standard: Learning Designs	Integrates theories, research, and models of human learning to achieve its intended outcomes.
Standard: Implementation	Applies research on change and sustains support for implementation of professional learning for long-term change.
Standard: Outcomes	Aligns its outcomes with educator performance and student curriculum standards.

Figure 1 Learning Forward and Delaware Standards for Professional Learning

In addition to outlining the characteristics or design of effective activities, these standards also show the importance of context. Conditions such as community, leadership, resources etc. affect whether or not professional learning promotes change. This will become a central idea in this ELP.

These definitions and standards for professional learning derive from a large body of research, which demonstrates the *potential* of professional learning to improve instruction and student achievement, but also that too often it falls short of these goals.

Problem Statement

As described above, professional learning for educators is ubiquitous as an improvement strategy. Examples of successful reforms highlight the critical role it can play. As Guskey (2000) states, “one constant finding in the research literature is that notable improvements in education almost never take place *in the absence of*

professional development” (p. 4). This is also a huge industry, with a federal investment estimated at \$2.5 billion per year (Layton, 2015). In a 2015 study, TNTP found that three large districts poured nearly \$18,000 *per teacher per year* into professional learning. Yet the return on all this investment is paltry and inconsistent. PD is far from a miracle cure or magic bullet. That is the conclusion of numerous research syntheses and policy studies (Borko, 2004; TNTP, 2015). Teachers concur – most give their experiences mixed if not critical reviews (Calvert, 2016; TNTP, 2015).

Research on Professional Learning

Since about 2000 there has been an effort to study the effectiveness of professional learning on student learning through large, rigorous, sometimes experimental studies, and to look more specifically at what design features of PD make the difference. These include meta-analyses (Blank and de la Alas, 2009; Yoon, Duncan, Lee, Scarloss & Sharpley, 2007), research syntheses (Desimone, 2009; Borko, 2004; Kennedy, 2016; Darling-Hammond, Hyler & Gardner, 2017) and large studies (Garet, Porter, Desimone, Birman & Yoon, 2001; Penuel, Fishman, Yamaguchi & Gallagher, 2007). One of the most commonly used frameworks comes from Desimone (2009) who posits the following five features:

- **Content focus** – Effective PD involves teachers in reinforcing academic concepts, then learning how to convey those concepts to students. So for example, in a math PD, teachers would do math and discuss pedagogical strategies for specific math concepts.
- **Active learning** – Effective PD engages teachers in a variety of activities that apply concepts to their practice. This could include hands-on activities or lesson planning time during workshops, or less

traditional PD experiences such as observing fellow teachers, being observed/coached, participating in school improvement or curriculum development processes. “Sit and get” workshops are the antithesis of effective PD.

- **Coherence** – Effective PD surfaces, and attempts to connect with, what teachers believe and know already (Desimone, 2009; National Research Council, 2000). It also aligns with the existing structures within which they work, such as state/national standards, curricula, and other concurrent reform initiatives. Learning is most powerful when embedded within the real contexts where teachers work.
- **Duration** – Teacher learning takes time. Of 16 programs identified that demonstrated significant gains in student achievement (Blank & de las Alas, 2009), the average total time was 91 hours and the average duration was 6 months. Desimone (2009) and Yoon et al (2007) suggest minimum durations of 20 and 14 hours respectively.
- **Collective participation** – Involving teams of educators from the same school or district has shown to be an effective design. It increases collaboration and peer support, makes it more likely that changes in instruction will “stick,” and creates momentum for change.

A recent framework (Darling-Hammond et al. (2017)) includes the same ideas and elaborates on them, stating that effective professional learning also uses models and modeling of effective practices, provides coaching and expert support, and offers opportunities for feedback and reflection.

These research-based features are practical and can be used to inform program design, delivery and/or evaluation. For example, I used them to create an observation protocol for the Title II evaluation (see Artifact E, Appendix E). Yet emerging research also cautions us not to apply them simplistically. In a review of 28 experimental studies, Kennedy (2016) argues that we need to look beyond the mere presence or absence of features. For example, conventional wisdom suggests that job-embedded supports like coaching and collaborative structures such as professional learning communities (PLCs) would yield strong results. Kennedy raises common-sense, yet easily overlooked, questions about quality and context. For example, how well trained are the coaches? What happens during coaching? Are PLCs engaged in meaningful or contrived work? What is the level of “buy in”? She concludes:

We need to replace our current conception of “good” PD as comprising a collection of particular design features with a conception that is based on more nuanced understanding of what teachers do, what motivates them, and how they learn and grow. We also need to reconceptualize teachers as people with their own motivations and interests. The differences shown here among PD methods of facilitating enactment strongly suggest the importance of intellectually engaging teachers with PD content, rather than simply presenting prescriptions or presenting bodies of knowledge (Kennedy, 2016, p. 30)

Similarly Timperley (2011) argues that more important than any particular design is teachers’ level of engagement in professional learning and what they do, or do not, learn from the experience. The ideas of teacher agency and engagement, as well as models of teacher learning, will be explored further later in this ELP.

Unfortunately, many of the professional learning opportunities available fall far short of these research recommendations, and do not fit the standards and definitions presented above. In an analysis of 376,908 activities in 203 districts and 27 states, Frontline Research & Learning Institute found that only 13% were sustained

over more than one meeting, only 9% occurred in an inherently collaborative format and only 8% were based on data about participants or their students. Workshops are the most common delivery format. Stand-alone, one-size-fits-all workshops still dominate (Combs and Silverman, 2016; Gates Foundation, 2014). Teachers have few opportunities to make choices about what they learn or to differentiate or personalize the activities to their specific teaching assignment, level of experience, or student needs (Calvert, 2016; Gates Foundation, 2014; Combs and Silverman, 2016). Perhaps most troubling, even when potentially research-aligned activities are in place, they may not be reaching their potential or meeting teachers' needs. For instance, many teachers report dissatisfaction with PLCs and lesson observations (Gates Foundation, 2014). These data all come from large, national data sets of teachers.

An analogous state survey provides insight into educators' perceptions in Delaware (New Teacher Center, 2013a). The Delaware Teaching, Empowering, Learning and Leading survey (TELL) was administered to licensed educators in districts and charters statewide in 2013.² The survey was online and anonymous with a response rate of 59%. Related results are shown in Figure 2, overleaf.

² The TELL Delaware survey was administered again in 2017 but the response rate was much lower: 4,030 respondents or 39.4%. Several high-need districts did not participate at all. Thus I use the 2013 data. Some improvements were seen in 2017 but I interpret those with caution due to response issues.

Q8.1 Please rate how strongly you agree or disagree with statements about professional development in your school.	
a. Sufficient resources are available for professional development in my school.	67.8%
b. An appropriate amount of time is provided for professional development.	66.9%
c. Professional development offerings are data driven.	76.0%
d. Professional learning opportunities are aligned with the school's improvement plan.	82.9%
e. Professional development is differentiated to meet the needs of individual teachers.	44.3%
f. Professional development deepens teachers' content knowledge.	60.2%
g. Teachers are encouraged to reflect on their own practice.	86.1%
h. In this school, follow up is provided from professional development.	55.8%
i. Professional development provides ongoing opportunities for teachers to work with colleagues to refine teaching practices.	64.5%
j. Professional development is evaluated and results are communicated to teachers.	42.4%
k. Professional development enhances teachers' ability to implement instructional strategies that meet diverse student learning needs.	71.2%
l. Professional development enhances teachers' abilities to improve student learning.	76.2%

Figure 2 TELL Delaware Data

While many of these responses seem relatively affirming, there are areas for concern. Differentiation of PD to individual needs and follow-up from PD both have relatively low levels of endorsement. Furthermore, disaggregated analyses indicated that respondent groups viewed professional development differently, and that teachers were more critical. For example, fully 98% of administrators agreed that professional development provides ongoing opportunities for collaboration but only 62% of teachers agreed. Ninety percent of administrators stated that follow up from professional development is provided, compared to just over half of teachers (54%) (New Teacher Center, 2013b). Why do administrators believe design features are in place but teachers do not experience them? These disparate data suggest the value of exploring stakeholder perspectives on professional learning more deeply, and attending to the complexities of implementation. They may also indicate the need to involve administrators more fully in professional learning.

As we have seen, the educational system in Delaware requires teachers to be learning all the time, developing new practices to meet changing student needs, standards, and assessments. Although clear state and national definitions and standards for professional learning exist and are derived from a large body of research, they are not yet being lived out the ground. The next chapter returns to the way universities can engage with these challenges. Specifically, based on my roles at UD, the chapter explores how research and program evaluation can help drive improvements in professional learning.

Chapter 3

IMPROVEMENT STRATEGIES

Theory of Change

The previous chapter introduced standards and research-based best practices for professional learning and demonstrated that we have far to go before we consistently reach them. This chapter argues that research and program evaluation about professional learning can contribute to improvement – and that universities have a crucial role to play. I recognize the challenge of developing strategies in one part of the educational system to effect change in another part. This improvement strategy is not simple but it is mutually beneficial for both higher education and the K-12 public schools, and it reflects my worldview. My career in education has spanned sectors and taken me from classroom teaching to program management and most recently program evaluation and educational research. Because of these varied experiences, I am oriented towards university contributions that make a real difference on the ground for teachers and students. Through my experiences as a participant, developer, facilitator, and evaluator of professional learning initiatives, I have become convinced that research can suggest powerful ways to improve them.

As sites of engaged and applied research, universities like UD can use several levers to effect positive change in professional learning. They can directly conduct research studies or evaluations of programs, and innovate with new methods of doing so. They can synthesize and disseminate findings from their own or others' research about professional learning. They can educate others in formal and informal ways about research-based professional learning. Finally, to make their research and

evaluations useful to K-12 public educators, they can continually learn about what those users value and need. Together, these activities can yield more informed professional learning leaders and better research and evaluation of professional learning. As a result, we expect to see stronger, more research-aligned designs, implementation and leadership and thus improvements in instruction and ultimately student learning. Figure 3 illustrates my overall theory of change.

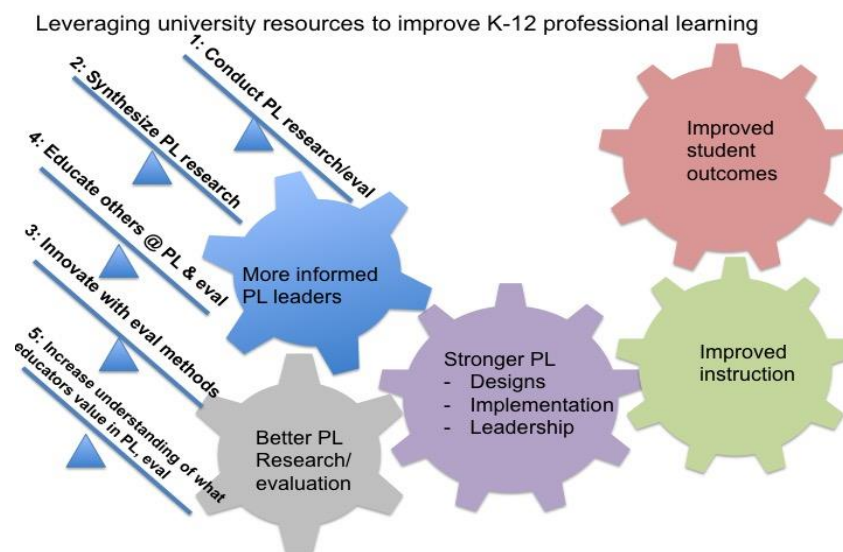


Figure 3 ELP theory of change

This theory of change is informed by three large bodies of literature related to how (a) universities can productively engage with public education (b) research and, in particular, evaluation can be used for program improvement and (c) professional learning can influence teacher instruction and student achievement. Mapped on to the theory of change, at the risk of oversimplifying, (a) corresponds to the five levers or improvement strategies (far left) (b) corresponds to the first three gears (blue, grey and purple) and (c) corresponds to the rest of the model. Literature about professional learning is discussed in Chapter 2. Literature about evaluation use (especially as it relates to professional learning) is introduced in the next part of this chapter and investigated more fully in Artifact K (Appendix K).

To learn more about university engagement with public education, I first situated myself within the University of Delaware (see Chapter 2). Considered more broadly, university - K12 collaborations bring together “the three major domains of schooling (i.e., curriculum and instruction, educational leadership, and student support) with the three functions of universities (i.e., teaching, research and service)” (Walsh & Backe, 2013, p. 596). Reading the broad literatures described above, I learned about the opportunities and also the pitfalls that can come when universities and public schools work together (e.g., Furco, 2013; Walsh & Backe, 2013). Such collaborations can mean many things, for many purposes, and a full review was beyond my scope. For this ELP, I focused on the domain of professional learning for in-service teachers (i.e., how those already in the classroom improve their instruction). I also focused on university functions related to my own position and responsibilities. I searched the literature and did not find exactly what I was looking for: a framework for how university-based evaluators can influence professional learning in K-12 schools. Instead, I turned to my practical experience of doing this work. The five levers represent ways that I understood and categorized improvement strategies that were within my span of control. As an external evaluator, most of my experience was with grant-funded professional learning programs. Therefore, I recognize that this work does not capture *all* the many, dynamic ways in which teachers learn, nor all the ways in which university personnel may participate in or encourage that learning.

Before examining the levers in more detail, I now take a step back to discuss evaluation of professional learning and why it is a powerful improvement strategy to ultimately move the “gears” of instruction and ultimately student learning.

Evaluation of Professional Learning

A central assumption of this ELP is that more rigorous, meaningful data about professional learning can be leveraged into program improvements that benefit teachers and, in turn, students. As a leading scholar in professional development evaluation, Guskey (2000) puts it:

A lot of good things are done in the name of professional development. But so are a lot of rotten things. What educators haven't done is provide evidence to document the difference between the two. Evaluation provides the key to making that distinction. (p. 94)

A brief discussion of terminology is necessary here. This ELP focuses on evaluations of professional learning programs but also uses broader research to inform my understanding. Distinguishing between program evaluation and research is complex and contested (Rogers, 2014). A commonly accepted difference is the purpose for which each is conducted. Evaluation focuses on use (Weiss, 1998). It answers questions, usually developed by, or in conjunction with, program staff. It can be used formatively (to adjust the program as it is being implemented) and/or summatively (to learn from its outcomes next time a program is being designed). By contrast, research focuses on generating knowledge, testing hypotheses or understanding phenomena. What we learn from research informs program evaluation. For example, research about professional learning (discussed above) helped me design evaluations and data collection instruments and is the foundation for standards that I used as criteria for evaluation. Thus, the two are not entirely distinct. Most of my artifacts integrate program evaluation and research in some way.

In the rest of this section, I review frameworks for the evaluation of professional learning, and outline ways to enhance current practice. Later artifacts

build on these ideas. There is a widespread call for more rigor and sophistication in evaluation of professional learning (Borko, 2004; Desimone, 2009; Kennedy, 2016). Of 1,300 studies reviewed by Yoon and colleagues (2007), only nine met What Works Clearinghouse evidence standards. Specifically, there is a need for more attention to impacts on student achievement. This push is reflected in the evaluation criteria for federal and state funding. For example, the application for Delaware’s Reimagining Professional Learning grant states that proposals must specify “systems for gathering and analyzing evidence of impact of professional learning on teacher practice AND student learning outcomes” (Delaware Department of Education, 2016a).

Current practice in evaluating professional learning programs often falls far short of these expectations. Indeed, the TELL data indicates that evaluation is a weakness in Delaware. Just 42.2% of respondents agreed that “professional development is evaluated and the results are communicated to teachers” (New Teacher Center, 2013a). When evaluation does occur, it is often limited to surveying participants to see if they were satisfied. A survey administered at the end of a program experience tells us little about what happens next. There are many steps between a satisfied participant and improvements to instruction, let alone student achievement. To dig deeper, it is necessary to flesh out how the professional learning activity is supposed to work (i.e., its logic model) and then to inquire into and measure what is actually happening for teachers and the students with whom they work. Since we know teacher learning takes time, so does evaluating or researching its impacts.

Frameworks for Evaluating Professional Learning: Guskey and Others

Evaluation frameworks present a systematic way to understand the intended effects of professional learning. In Delaware and nationally, the most prevalent

framework is Guskey (2000), which states that evaluation should attend to five levels of effects: participants' reactions, participants' learning, organization support and change, participants' use of new knowledge and skills, and student learning outcomes. Each level is a pre-requisite for the others yet success at one level does not guarantee success at the next. For example, a participant can love a professional learning experience and learn a lot in it (Levels 1 and 2), but if his/her context fails to support the new learning (Level 3), s/he may never demonstrate changes in teaching practice (Level 4) and student outcomes may not improve (Level 5). This could happen if structures such as the school schedule, teaching/PLC arrangements, or curriculum requirements compromise the implementation of professional learning, or if school leaders or fellow teachers did not understand or support it. (This is an issue of coherence). To understand the intricacies of professional learning transfer, it is necessarily to evaluate all five levels. As the levels increase, evaluation becomes more complicated and time-consuming and yet the information gathered is more valuable. Guskey argues that even managers of small programs can inquire into all five levels, and external evaluators certainly should.

The Guskey framework has the advantage of being familiar, straightforward, and endorsed within the Delaware educational system. For example, proposals to DDoE's Reimagining Professional Learning had to anchor their evaluation plans in this framework, and grantees reported program evidence by what Guskey level it demonstrated ((Bennett, 2017). For these reasons, I based most of the artifacts in this ELP on Guskey (2000). Three implications of this framework are particularly relevant for my work. First, evaluating level 3 ("organization support and change") involves looking at school or district conditions and culture and at how these influence the

implementation of professional learning activities. Here Guskey's framework fits well with Kennedy's (2016) argument about the importance of context. Educational leaders and other practitioners may benefit from an outside perspective to accurately perceive and evaluate their context. Second, completing a full Guskey-aligned evaluation requires a substantial timeframe and budget. Teacher learning is slow and messy, and longer-term studies capture greater changes in instructional practice and student learning (Kennedy, 2016) yet funders may impose unrealistic evaluation timelines. This was sometimes a concrete challenge in my work, especially in one-year projects. Finally, although the Guskey framework is user-friendly, practitioners may need support to translate it into practice. They also may require additional resources or capacities to fully conduct an evaluation. These three implications are also opportunities for universities to contribute to the evaluation process.

As mentioned above, another contribution universities can make is to synthesize ideas from literature. With this in mind, I also tried in this ELP to move beyond Guskey to consider other frameworks for professional learning or models of how teachers learn. My annotated bibliography (Artifact C, Appendix C) reflects these efforts. I undertook this exploration partly out of my own curiosity and partly in response to concerns and questions raised from various directions including educational leaders with whom I worked. Some noticed that Guskey's five-level framework did not apply to their situations or did not adequately explain why expected changes in teaching or learning were – or more often, were not – occurring. Through my reading I learned that other models for professional learning do exist.

All ten of the other models I identified cite Guskey; they often explicitly negotiate with his ideas. There are four major differences. First, some aim to build

onto the Guskey framework, for example by categorizing types of learning activities or elaborating on what kinds are most likely to propel change (Desimone, 2009; Kennedy, 2014; Kennedy, 2016). Second, others put more emphasis on individual teachers and their motivations and other antecedents of change, thus pushing back on the idea of a generic model of professional learning (Clarke and Hollingsworth, 2002; Opfer and Pedder, 2011; Evans, 2014). Although the teacher is the key actor in the Guskey model, some authors argue that his/her agency and engagement is under-explored (Boylan et al., 2017). Third, some models push more deeply into what Guskey would call “Level 3,” looking at the role of school leaders (Evans, 2014; Timperley, 2011) and other aspects of organizational or political context within which professional learning takes place (Kennedy, 2014; Opfer and Pedder, 2011). Fourth, some models fundamentally disagree with the linear and “leveled” nature of the Guskey framework (Coldwell and Simkins, 2011). Instead they conceptualize professional learning as multi-directional process (Clarke & Hollingsworth, 2002), a cycle (Timperley, 2011) or a complex, nested system (Opfer and Pedder, 2011).

I have not had the opportunity to base an evaluation on one of the alternative models – nor, given the continuing prevalence of Guskey in Delaware, does it seem wise to depart from that framework entirely. Still, I agree with Boylan et al. (2017) that researchers and practitioners can benefit from knowing multiple models and can combine insights from more than one of them. My most recent research project, Study of the Teacher Leadership Initiative (Artifact B, Appendix B) incorporates some insights from what I learned as I explored models of teacher learning. For example, I collected more data than I previously had about teachers’ motivations to participate.

Five Levers for Improvement

As depicted in Figure 3, the overall goal of this ELP is to use the research, evaluation and teaching capacity of universities to educate professional learning leaders and to improve the quality of research in the field. Specifically, I identified five ways of doing that, or levers. For each lever, I developed one to three artifacts that demonstrate how I used it.

Conduct research and evaluation related to professional learning.

This is the most straightforward way university staff can contribute. Earlier, I described the purpose of research as generating knowledge and the purpose of evaluation as program improvement. University professionals can be involved in either. My ELP includes one artifact that is an external evaluation of a grant-funded program and another that I would consider research, since its purpose was to learn from and understand the dynamics of a professional learning process. Both artifacts demonstrate how a university contributes human, technical and intellectual resources.

In 2015, I and three DERDC graduate students conducted an evaluation of the activities of Delaware's Title II grant (Artifact A, Appendix A). This grant provided professional development in high-need schools and districts statewide and comprised seven distinct activities. Although similar in some ways, these activities varied in scope, both the number of participants and the span of time. In conducting this evaluation, we tried to balance consistency and particularity. We asked the same evaluation questions for each activity and went into more depth where data were available to support our inquiry. The questions drew on Guskey's levels of professional development evaluation: (1) Did the project reach the target audience? (2) Was the PD provided high quality? Were participants satisfied? (3) Did teachers

develop the target content knowledge and pedagogical content knowledge? (4) Does teachers' instruction demonstrate the target skills/knowledge? And (5) Does student achievement in the target skills/knowledge improve? This report demonstrates how I anchored my evaluation in the Guskey framework and collected data through surveys, interviews, observations, and content analysis. It also demonstrates how evaluation makes recommendations to strengthen professional learning. For example, we recommended that in the future Title II increase coordination among parts of the project and establish shared parameters for PD design.

The second example of this lever (Artifact B, Appendix B) is a study of the Teacher Leadership Initiative (TLI), another statewide program run by a non-profit called Schools That Lead (STL). This was more exploratory research than contracted evaluation. STL approached DERDC about the feasibility of a research partnership, and we worked together to design a process for qualitatively documenting the early impacts of the TLI on participating teachers, their principals, and their schools. TLI is essentially a teacher inquiry project in which participants pose questions about their students' learning, then work with a peer or a small group to collect data, reflect on it, and shift their instruction in response.

The study engaged six questions: (1) Who gets involved in TLI, how, and why? (2) How do participants describe their experiences within TLI? (3) How do relationships between TLI teachers, other peers, and administrators develop during the TLI? (4) How does participating in TLI influence the instructional practice of participants? Others? (5) What impacts are evident on student learning? (6) What organizational conditions and supports facilitate or constrain the work of TLI participants? The primary mode of data collection was interviews (17 teachers, 10

principals or assistant principals), along with some observations and review of project materials and records.

Although the Title II evaluation is more typical of my work at DERDC, the TLI study demonstrates my attempt to use some new insights and in-depth qualitative methods to understand teacher learning. Rather than providing recommendations per se, it concludes by highlighting three tensions within the TLI model. These were the relationships that formed among the small group of early adopters vs. the need for the project to expand its reach; the importance of school administrators actively supporting TLI yet remaining “hands off”; and the challenge of aligning a teacher-directed professional learning activity with other improvement efforts.

Synthesize research related to professional learning; develop related recommendations.

Practitioners want their efforts to be guided by research but often lack the time or capacity to stay abreast with it. One contribution university staff can make is to read literature so that educational leaders do not have to, then distill it in pragmatic ways. They can also condense and disseminate ideas from their own research. My ELP shows three different examples of this strategy. First, as described above, I created an annotated bibliography of ten frameworks for professional learning, identifying implications for practitioners or other researchers, and comparing each to the Guskey (2000) framework (Artifact C, Appendix C). I did this to develop my own understanding of how teachers learn, to become a more sophisticated evaluator, and to look for usable insights for others. I used academic search techniques to locate articles (e.g., following citations and reference lists) and in the process discovered some new journals. I found that looking outside of the United States (even at other Anglophone countries) brought a valuable new perspective on teacher learning. In particular,

models from less accountability-driven educational systems seemed to pay more attention to teachers as individuals, with backgrounds, interests and needs of their own, rather than simply as producers of student achievement results.

Second, I wrote a practice brief about improving professional learning in Delaware (Artifact D, Appendix D). This summarizes research discussed in Chapter 2 here and includes vignettes of four exemplary programs in the state: the TLI, the Delaware Reading & Writing Project, Partners4CS, and the Laurel School District Comprehensive Professional Development Partnership. I had evaluated all but the Laurel program during my time at DERDC, and this brief gave me an opportunity to share my insights from program evaluation more broadly. The audience for the brief includes teachers, administrators, policymakers, and university partners. Through the vignettes, I wanted to illustrate what research-based programs look like in action and also to draw four lessons about professional learning. These include the importance of teacher agency; the balance between internal and external sources of expertise; the potential of reciprocal partnerships and of professional learning to influence policy; and the power of program coherence and administrative support.

Building on this final idea, I created a presentation to highlight the important role school leaders play in implementation, “After the PD.” I chose this topic for a literature review and presentation in EDUC 890, since I felt leaders needed to better understand their influence over professional learning. This turned into a narrated PowerPoint targeted at school leaders and those who support them such as district-level administrators or leadership coaches (Artifact E, Appendix E). It summarizes research about how school leaders can support teachers’ putting new skills and knowledge into instructional practice. The presentation uses the five roles of the Kose

(2009) framework (visionary leadership, learning leadership, structural leadership, cultural leadership, political leadership); provides concrete examples of each role from around the state and the country; and guides leaders in thinking through implications and action steps for their work.

Innovate with evaluation methods.

Universities also generate new ideas and knowledge. In my field, this entails creating innovative methods for evaluating professional learning given the shortcomings often noted. I have limited experience using this lever but I included one example (Artifact F, Appendix F), a meta-evaluation of a 2013-14 DDoE competition, the Specific and Innovative Improvement Practices (SIIP) grant. SIIP funded 14 projects focused on the same set of general goals (one or more of the following: (a) teacher-led projects that drive improved student outcomes, (b) Common Core implementation and assessment, (c) student supports and dramatically improved school climate or (d) accelerating the achievement of underperforming groups) but ranging widely in scope and strategy. Almost all included professional learning.

Our innovation was to create a method to synthesize across these 14 very different projects to draw overall lessons. We designed evaluation rubrics for implementation and outcomes of the SIIP projects and conducted meta-evaluations using the 14 SIIP project reports as our data source. To my understanding, this was the first use of evaluation rubrics in Delaware. This was a collaborative project. The DERDC Director led the work and conceptualized the method. Together with a graduate assistant, I drafted the rubrics. There were 16 rubrics in three sets: process (7 rubrics), outcomes (5 rubrics) and goal-specific rubrics (4). As a team of three we analyzed and interpreted data and wrote the report. The rubric structure allowed us to

analyze data by project (i.e., to calculate total and subtotal scores for each of the 14 funded projects) or by rubric (i.e., to look across the projects to see how they performed in a given area). In the first analysis, this allowed us to examine reasons associated with project performance (e.g., design of activities, realistic scope, district resources and support). In the second analysis, we found that process rubric scores were generally higher than outcome rubrics and that the lowest average rubric score was in Application of Professional Development.

Another aspect of this project was providing technical evaluation assistance. We met once with each project early in the grant year and several times with DDoE and made ourselves available for further support. We aimed to help local projects and the state develop their own evaluation capacity.

Educate others about research-based professional learning design, implementation, and evaluation.

This lever involves using a university's teaching mission to improve professional learning. I used this lever both formally and informally as an instructor in the Masters in Teacher leadership program, as a mentor for graduate research assistants, and as a student myself. This lever influences those who are, or will be, professional learning decision-makers to understand research in the field. In one case, I contributed to a new course about professional learning, EDUC 774, part of the Masters in Teacher Leadership core curriculum. I was the preceptor for the instructor, Dr. Chrystalla Mouza, in Spring 2016, the first time this course was offered. The course prepares current and aspiring teacher leaders to identify, design, lead, and evaluate professional learning programs in their school or other workplace. My role included reviewing and providing feedback on the initial course design; responding to

student questions; facilitating discussion boards; monitoring student progress and grading assignments, including the Professional Development Plans (PDPs) that our students developed. Artifact G (Appendix G) includes my reflections on teaching this course and on student progress; it also summarizes the course evaluations we received.

My other two artifacts for this lever show how I applied EdD coursework to my job in program evaluation. In both cases, I created new opportunities to engage others in learning about professional development. As a supervisor of graduate research assistants at DERDC, part of my job was preparing them to contribute as members of evaluation teams. Especially for new students, this required building their understanding of professional learning as well as ways to evaluate it. In EDUC 897, our final project was to develop a curriculum module. To make that authentic, I developed a module based on the Understanding by Design (Wiggins and McTighe, 2006) format around the essential question: What will it look like when professional development is successful? (Artifact H, Appendix H). As Wiggins and McTighe recommend, I planned backwards from a “desired real-world application”: for GRAs to be able to evaluate professional learning. When I designed this module (2014), DERDC was anticipating a 100% turnover in graduate research assistants and building Center capacity was a crucial concern. The unit builds up to two performance tasks and includes a sequence of reading/learning and field experience activities.

Artifact I (Appendix I) demonstrates teaching in another, job-embedded context. This was my final project for EDUC 818, Educational Technology Foundations. It is an online collection of resources and best practices to help people plan, design, implement, evaluate, or select high-quality online professional development (OPD). I became interested in OPD when two clients began offering

online or blended opportunities. It was clear that while some principles of face-to-face experiences translated, the online environment posed new for professional developers and evaluators and I saw an opportunity to provide research-based guidance. I gathered resources through researching the literature and best practices of OPD. My toolkit draws on the National Research Council (2000) framework (i.e., that learning experiences should be knowledge centered, learner centered, assessment centered, and community centered). It includes examples of effective OPD designs; advantages and obstacles to OPD; implementation considerations including resources, leadership, and accessibility; and summaries of research on OPD.

Increase understanding of what educators value in professional learning efforts and evaluations.

Although perhaps abstract, this strategy undergirds my understanding of how researchers or evaluators and practitioners should collaborate. In describing levers for improvement originating from a university environment but used within the K-12 sector, one risk is implying that schools need universities or are blank slates, that they lack their own expertise or interests. This is not my intent at all. As a university employee with roots in K-12 education, I was discouraged when colleagues failed to grasp to the real world of schools or the professionalism of those who worked in them. This lever is about establishing two-way communication channels and ensuring that university-based staff learn from practitioners and listen to what they need and value. Artifact J (Appendix J) examines what teacher leaders and other students from EDUC 774 wanted out of professional learning. It does this by analyzing the professional development plans (PDPs) they submitted as their final projects for the class. After obtaining consent from former students, I analyzed a set of 33 PDPs written for

elementary, middle or high schools in Delaware. I reviewed them with a framework of ten features, developed from three overlapping sources: the Standards for Professional Learning, the Desimone (2009) research framework, and the syllabus for 774. First, I determined whether each PDP contained clear and specific, some or partial, or no evidence of each feature. Second, I analyzed the data qualitatively for themes.

Artifact K (Appendix K) turns the lens to professional learning evaluation and interrogates the central assumption of this ELP, that professional learning evaluation can lead to program improvement. I interviewed four professional learning leaders and former clients, seeking feedback about the evaluations I had conducted of their programs while at DERDC. My interview protocol was grounded in the literature about evaluation use, the Utility standards for program evaluation (Yarbrough, Shulha, Hopson & Caruthers, 2011) and instruments to assess the utility of an evaluation (e.g., Stufflebeam, 1999) or to obtain feedback from clients (e.g., Doino-Ingersoll, Haley, Dowell & Chambliss, 2005). My report describes the four programs and their evaluations; explores client understanding and expectations about evaluation; examines whether and how the evaluations were used as well as facilitating or constraining factors for use; shares client feedback, both positive and negative, on the evaluation and my work; and reflects on what I learned through the inquiry.

Chapter 4

IMPROVEMENT STRATEGIES RESULTS

This chapter analyzes the results of each improvement strategy. As my theory of change indicates, these strategies are aimed at two primary outcomes: (1) informing professional learning leaders and/or (2) improving professional learning research or evaluation. For each strategy, or lever, I looked for evidence that I reached either or both these outcomes. In a few cases, I also found evidence of a subsequent outcome: stronger professional learning designs, implementations, or leadership. By informing leaders and/or improving the quality of professional learning research or evaluation, I influenced the professional learning programs. Not all strategies were successful; where results were not achieved I attempt to explain why. The next chapter synthesizes across and reflects on these results more deeply.

Conduct research and evaluation related to professional learning.

My ELP included two studies I conducted of professional learning programs (Artifacts A and B, Appendices A and B). For both the Study of the Teacher Leadership Initiative and the Title II evaluation, I succeeded in informing professional learning leaders. When I shared the TLI study with Schools That Lead (STL) program staff. I received positive feedback. STL stated that it was “solid” and “fun to read,” that it accurately captured their program and they would like to work together in the future. In our debrief conversation, it was clear that the report confirmed impressions that STL had already formed about their program and directions in which they were already planning to develop it, rather than generating brand new ahas! For instance, I

found that teachers had mixed feelings and experiences about scaling the inquiry process beyond the initial group of TLI participants. My client wrote that it was “not too surprising to read some of the lack of knowledge on how to scale and struggle to scale” and stated that the program had increased supports for this process in subsequent cohorts. I would argue that program staff can become “more informed” about professional learning even if the findings are confirmatory rather than revelatory. For example, they may develop more confidence in their interpretations of what is happening or they may expand their understanding of stakeholder perspectives. Some findings were new information for STL staff; these related mostly to organization and structure (e.g., records management) than to deeper professional learning impacts. STL disseminated key findings from our study through their monthly newsletter (Schools That Lead, 2017) thus broadening the scope of “more informed professional learning leaders.”

To understand the outcomes of my program evaluations, I conducted a feedback study with four former clients who led professional learning activities including two from Title II. This artifact focused on evaluation use (Artifact K, Appendix K). I wanted to understand whether and how my evaluations provided information, and in what ways that information had been used for program improvement, decision-making, or any other purpose. In my theory of improvement, evaluation use is the mechanism by which “informing professional learning leaders” turns into “better professional learning designs, implementations and leadership.” Through my interview protocol (described in Chapter 3) I asked about three kinds of use: instrumental, conceptual, and process. Instrumental uses are concrete and involve action and/or decision-making while conceptual uses are broader, involving changes in

thinking. In process use, participants learn by being involved in the evaluation rather than from its findings. I learned that all four evaluations were indeed used, though in one case the use was quite limited. Instrumental uses were the most common.

I learned that my work, including the Title II evaluation, informed professional learning leaders in a variety of ways. At the most basic level, they provided information necessary for accountability (e.g., reporting to funders). More significantly, they guided decisions related to program scope, budget, schedule, participant selection/admission, or staff management. For example, in one program evaluation data demonstrated the need to focus activities. In another, the evaluation gathered information about what partnering schools/districts wanted to see from the program. This in turn affected program recruitment and the design of activities. Every client I interviewed could give at least one example of a time they used specific information from my evaluation for program improvement. One surprise was learning what types of information clients most valued. More than official evaluation reports, they particularly appreciated informal and ongoing communication with me, such as the questions I asked as we debriefed or planned a program activity. My interpersonal and communications skills were critical to using this lever for improvement.

Program context and conditions affected use. I gathered data from two multi-year and two one-year projects, and found more use in the longer projects, because relationships, trust, and understanding of the role and utility of evaluation all develop over time. I also found the dynamics of use were more straightforward in smaller projects with flatter organizational structures, where I could establish direct relationships with the staff members who would be acting upon evaluation insights. One project reflected in Artifact K was larger (\$5M, with four PIs and multiple work

groups). The task of informing others grows ever more complex the more “leaders” are involved in the professional learning, especially if they have different interests.

I identified several examples where my evaluations contributed to stronger professional learning design, implementation, or leadership. In one case, project staff were scientific experts but did not know much about adult or professional learning. Through data I collected and questions I asked, I helped them realize the need to engage participants more actively. In another project, the evaluation demonstrated that while participants greatly enjoyed a workshop, they did not know how to translate activities to their classroom settings. The next year, project staff refocused the activity to increase coherence. In a third project, the evaluation helped improve the quality of a research-supported design, instructional coaching. By collecting and analyzing data from multiple perspectives (participants, coaches, and school administrators), the evaluation identified weaknesses in the coaching model which the project then attempted to address through more resources and support.

Finally, there was some evidence that I contributed to improving the quality of professional learning research and evaluation. Two projects had research components: as we were conducting the evaluation, other investigators were conducting parallel research about professional learning. I contributed data and analysis and was listed as a co-author on several presentations and manuscripts. More broadly, clients with whom I worked long-term told me that they learned about the purpose and value of program evaluation and that they developed their own skills of data collection, program observation, or other forms of evaluative thinking. This could be considered process use. I also learned that multi-year projects allowed me to better “walk the talk”

and model high-quality professional learning evaluation while one-year timeframes were limiting.

Synthesize research related to professional learning; develop related recommendations.

This lever provides the opportunity to inform professional learning leaders but I have only limited evidence of success. My practice brief on Improving Professional Learning in Delaware (Artifact D, Appendix D) will be published by UD's Partnership for Public Education (PPE) in 2018, made available through the PPE website and perhaps disseminated through other channels. It provides valuable recommendations to teachers, administrators, system leaders and policymakers. I shared my presentation on After the PD: The Role of School Leaders in Implementation (Artifact E, Appendix E) with professionals at the Delaware Academy for School Leadership (DASL). They planned to suggest it as a resource for those working with and supporting school leaders, such as district-level instructional supervisors or for leadership coaches. Again, although the goal was to provide actionable insights from research, data are not yet available about whether and with whom it was used, or with what results.

In some ways, the primary audience for my annotated bibliography (Artifact C, Appendix C) ended up being myself. I gained new insights by reading more widely and theoretically about teacher learning, but it was challenging to develop pragmatic applications for others. This may be because of the genre or structure of the piece, or because I need more practice in synthesis. However, producing this artifact helped me grow as an evaluator and thus contributed to better research of professional learning. For example, I applied new insights about teacher motivation and agency in the TLI study (Artifact B, Appendix B), in which they emerged as major themes.

Innovate with evaluation methods.

The goal of this strategy is mostly to improve research and evaluation of professional learning. In the SIIP project, my DERDC colleagues and I developed the meta-evaluation method so it could be used in the future. Indeed one of our recommendations in the report (Artifact F, Appendix F) was to “continue to develop and use the evaluation rubrics to evaluate DDoE initiatives.” To my knowledge, this has not occurred. We at DERDC certainly could have done more to share our innovation more broadly (e.g., present at conferences).

In the SIIP project, we also aimed to provide technical assistance to improve the quality of professional learning evaluation by grantees. Here too, our success was limited; five of the 14 projects requested and received assistance from DERDC (e.g., on evaluation instruments, data analysis, or program observations). Although we offered help on numerous occasions, there was no systematic incentive or push for projects from their sponsor (DDoE) to take advantage of it, and overall the evaluations were inconsistent. Therefore, we recommended that DDoE “set explicit expectations for not only professional development, but also for project monitoring and evaluation procedures and reporting of evaluation findings.” There is some evidence that subsequent DDoE grant programs included more explicit guidance and support, but that change cannot be linked in any way to our work. Recently, DDoE has shared insights and lessons learned through the Reimagining Professional Learning grant, discussed elsewhere in this ELP. These include that “LEAs need more support and models for evaluating professional development and data collection” (Bennett, 2017). In partnership with Learning Forward and WestEd, DDoE developed tools and systems to provide that support. These included protocols for calibrating evidence across programs; online learning opportunities about professional development and

evaluation; reporting templates aligned with the Guskey framework, etc. Ultimately it seems DDoE arrived at the destination we were recommending and took steps to improve the quality of professional learning evaluation, but in its own way and with non-university partners.

My DERDC colleagues and I received a positive response to SIIP report from our client. DDoE praised our method and stated that the information in the report was valuable. However, external events interfered with use of our recommendations. In the mid-spring of the grant year, funding cuts were made. Thus, by the time DDoE received our report, they already knew there would not be another SIIP competition. This likely reduced their investment in our SIIP-specific findings. When DDoE launched subsequent competitive grants for professional learning, different personnel were in charge and probably were not even aware of the work that had been done with SIIP and DERDC.

Educate others about research-based professional learning design, implementation, and evaluation.

I used this lever most effectively by helping Dr. Chrystalla Mouza launch EDUC 774, Designing Professional Development. By several measures, this new course successfully made 58 current and future teacher leaders more informed and effective with professional learning. This is evident in both Artifact G (Appendix G), my reflections on the course and our evaluation data, and Artifact J (Appendix J), an analysis of student work products. EDUC 774 had a large enrollment (58) and positive learning outcomes: all 58 students completed all requirements, and 71.7% earned a grade of A. Student feedback on the course was enthusiastic. By large margins, students found EDUC 774 relevant (93.5% agreement) and well organized (96.8%

agreement) and were overall satisfied with the course (96.8% agreement). As explored in Artifact J and described more fully below, the student PDPs show strong understanding and application of many course concepts. Since students left our class with a complete plan, they at least had the potential to create stronger professional learning designs.

I also claim success in being invited to act as the instructor for EDUC 774 in Spring 2018 – and before that, to serve as a preceptor for three additional classes in the Masters in Teacher Leadership program (Fall 2016, Spring 2017, Fall 2017). I view this as an endorsement of my teaching skills and a chance to continue to develop them. For example, I have made revisions to EDUC 774 based on areas where our students needed more support, scaffolding and information in 2016.

Turning now to DERDC's graduate students, I partially achieved the goals of building organizational capacity to conduct high-quality evaluations of professional learning and informing students about the field. However, I reached this goal in a different way than originally intended. For two reasons, I did not implement my curriculum unit about professional learning and its evaluation (Artifact H, Appendix H) fully as written. First, DERDC unexpectedly brought on one student for the summer (2014). Since we had a compressed timeframe for orientation and a single new student rather than a cohort, I could not use the full curriculum. Luckily, this individual learned very quickly and was ready to go "into the field" even without all the preparatory experiences outlined in the curriculum. He and another (veteran) GRA became co-authors on the 2014 Title II report and contributed in meaningful ways. They did ultimately increase their understanding and skill in program evaluation. Later, DERDC welcome two more GRAs in fall 2014 and experienced the opposite

situation. The students had more basic developmental needs that dominated our time when they first arrived. Again, real-world events interfered with the delivery of the curriculum. Still, I drew from it what I could (e.g., teaching and using the Guskey framework, reflecting on our own PD experiences, co-observing program events) and both students did eventually get to the point of being able to independently collect, analyze, and report professional learning evaluation data.

I did not publicize my online professional development (OPD) toolkit (Artifact I, Appendix I) or teach it in any official way, but I drew on what I learned from it to design data collection instruments and more informally to advise project staff.

Increase understanding of what educators value in professional learning efforts and evaluations.

This lever is the most conceptual of the five. It is important for university staff to listen to what practitioners want and value in their professional learning and/or program evaluations. Doing so increases the chances that professional learning programs will meet participant needs and that the evaluations of these of these programs will be used. As I worked on Artifact K, I realized that I had rarely formally asked evaluation clients for feedback. I developed an effective interview protocol and discovered that former clients gave constructive criticism, which in turn improved the quality of my professional learning evaluation going forward. For instance, I learned that I need to do more to establish relationships with project staff early in the project, especially if they are skeptical or anxious about evaluation. I gained ideas for building a shared understanding of the evaluation (e.g., provide a model evaluation report so clients know what to expect). Although initially intimidating, these interviews turned

out to be richly rewarding. They showed me the value of structured two-way communication channels between evaluators and professional learning leaders.

My analysis of professional development plans (PDPs) from EDUC 774 (Artifact J, Appendix J) was another attempt to listen to educators. PDPs provide insight into how a group of current and aspiring teacher leaders envision professional learning that meets their, and their schools' needs. My findings confirm that teachers want professional learning to be relevant, collaborative, and personalized. Relevance was generally achieved by rooting the activity as "close to home" as possible, such as within the same content area or grade level. All 33 PDPs (100%) incorporated collective participation by groups of colleagues and mentioned the importance of differentiation, though only 24 (72.7%) provided details about how this would be achieved. The plans demonstrate a strong interest in leveraging teacher expertise from within the school. Many PDPs called upon teachers to model practices, facilitate sessions, coach their peers, or oversee the entire effort.

My analysis also revealed some areas of relative weakness, features with which the PDPs were less aligned. Two stand out. Less than half of PDPs (42.4%) included clear and specific data demonstrating a need for the planned activities. Just over half (54.5%) described detailed plans for leading/facilitating the activity. Moreover, while students generally gravitated towards using "in house" facilitators, they rarely addressed capacity or resource issues. These data may suggest that educators do not value these professional learning features as highly as some of the others. Alternatively, they may indicate that our students had not thought as much about them. This is likely given the wide range of professional and educational experience in the class. Some were already teacher leaders while others were in the first few years of

teaching. Some were finishing the masters' program while others had just begun. The PDPs in general were somewhat over-optimistic about teachers' willingness and ability to adopt new practices, and about the resources required to successfully implement professional learning initiatives. This analysis helped me identify needed changes to EDUC 774 for Spring 2018. For example, I revised the project template to put more emphasis on data and leadership, and to add an assignment in which students present their PDPs to their supervisor or other stakeholders. I anticipate these changes will result in even stronger PDPs this year.

Chapter 5

REFLECTION ON IMPROVEMENT STRATEGIES RESULTS

Now I reflect further on the specific results described in Chapter 4, highlighting seven lessons that I learned through this work. At the end of this chapter, the lessons are illustrated on a revised Theory of Change diagram (Figure 4).

The first lesson was reinforcement of a qualitative concept: the researcher or evaluator's positionality matters. This is my story of using particular improvement strategies to improve professional learning, and it reflects my skills, interests, and areas for growth. A different person might achieve different results with the same levers; I can think of colleagues for whom this would be the case. Therefore, it is difficult to generalize about what improvement strategies from this ELP should be continued, redesigned, or dropped. Rather, university professionals can influence K-12 professional learning in a variety of ways; the key is aligning staff members' skills and passions with the improvement strategies they use.

Second, I found that some strategies have more leverage. Overall, I was successful in informing professional learning leaders. As described in the last chapter, I accomplished this in a variety of ways, for a variety of audiences. I discovered that the most powerful levers were directly conducting research and evaluation of professional learning (Lever 1) or teaching others about it (Lever 4). As a researcher or evaluator, one's influence on professional learning is always indirect, mediated through other people. Still, some of my levers are more powerful than others. I seemed to add the most value when I was involved in more hands-on ways, and where I could use my communication and interpersonal skills to develop relationships with

professional learning leaders and participants. This also meant that smaller projects were a more conducive setting for my strengths.

My third lesson was about how I personally could make professional learning research and evaluation “better.” The consensus is that scholarship in this field is superficial and simplistic and does not do enough to trace the impacts of professional learning on instruction and achievement. I recognized, and experienced, some real-world reasons why this is the case. The primary culprit in my experience was short or constrained funding parameters for both programs and their evaluations. Based on my skills and qualitative orientation, I wanted to go deeper in understanding and communicating the experiences of those involved in professional learning. I also wanted to develop a more sophisticated model of how that learning occurs. I started to see more limitations to the Guskey framework: it is linear and does not sufficiently account for individual learners and their motivations, nor for the characteristics of powerful learning experiences. Within the resources available in evaluations, I tried to do more to capture the “messiness” of teacher learning, but it was not until the TLI study that I could really apply some of these ideas. Because TLI was an exploratory study and not a commissioned evaluation, I had more freedom in designing the questions and the methods. Instead of asking, “did the program work?” I could ask, “for whom did it work, under what circumstances, and why? Also, what does it mean to “work,” and who gets to decide?”

Fourth, I learned that given my predilection for big questions like that, developing my skills of synthesis is key. This portfolio included three artifacts demonstrating the strategy: *synthesize research about professional learning, develop related recommendations*. Each one challenged me greatly. Although I never found the

magic “trick,” I got better at negotiating clarity and complexity. Sometimes, that meant seeking committee feedback and revising artifacts to achieve a better balance. This lever is more effective when the audience is clearly defined and when the researcher/evaluator has some entrée with them. Given the breadth of my roles, this was not always the case for me.

Fifth. I learned about research or, particularly, evaluation use. For the gears in my model to turn, evaluation must yield relevant information which leaders must use to spur improvements. The follow-up studies I conducted with former clients gave me insight into when this happened. I learned that timely, clear, and candid communication facilitated use, and that clients were more likely to use insights from ongoing dialogue with me rather than formal reports. I learned that my clients, like me, found evaluation “better” when it went deeper into changes in instruction and learning, and when it accessed perspectives that they may not have had the time, ability, or awareness to reach on their own. The literature suggests that participatory methods promote evaluation use. However, I found that was only somewhat true. Although I am fascinated with non-traditional methods that collapse the “evaluator” and “practitioner” boundaries, and would love someday to innovate with them, these methods require more time than my clients had.

This leads to the sixth lesson: just like professional learning, evaluation is situated. External conditions such as resources (time, funding), contract timelines, political environment, and leadership matter, as do internal organizational dynamics. My evaluations were used most in situations where there was time to build relationships and trust with clients and for them to discover the value of evaluation, and where leaders set the expectation of using feedback for continual improvement,

and where the evaluation aligned with broader organizational strategy. These facilitators are similar to those I discovered in professional learning. Indeed, I saw that my model really included application of new skills and knowledge at two distinct points: when professional learning leaders use evaluation, and when educators act upon the professional learning activity.

Of course, the conditions for professional learning or its evaluation are rarely optimal. For example, one-year contracts are common (as documented in the SIIP and Title II evaluations). Yet they clearly do not allow for the slow and messy process of teacher learning (nor program improvement). Professional learning leaders may know that this but be stuck with other constraints (e.g., federal timelines, budget cuts). This poses a dilemma: should university staff not get involved with programs if they cannot strongly align with the research? More relevant for me, should evaluators not bid on such contracts? My answer to this question, influenced by colleagues and supervisors, was “no.” Rather, I attempted to make the best of it, doing the highest-quality work possible under budget and time limits. Similarly, we tried to use research and best practices of professional learning as our standard while also being realistic about the constraints that our clients’ programs were under.

Finally, I recognized how pervasive leadership is in my theory of change. Most evaluations and studies I have conducted, including all the ones in this portfolio, demonstrate the importance of educational leaders before, during and especially after professional learning activities. This message was the point of my presentation in Artifact E (Appendix E). I learned how complex and nuanced it is for a leader to support teacher learning. As one of the TLI principals told me, it requires simultaneously being “hands off” to let teachers direct their own inquiry and make

mistakes, and yet “hands on” in providing resources and both logistical and political “cover” for the process to occur. In my work with teachers and educational leaders across a variety of settings, I will keep trying to communicate and encourage that balance.

Figure 4, below, revises my theory of change to include these seven lessons, labelled L1 – L7. Although now more complex, the core theory is still the same.

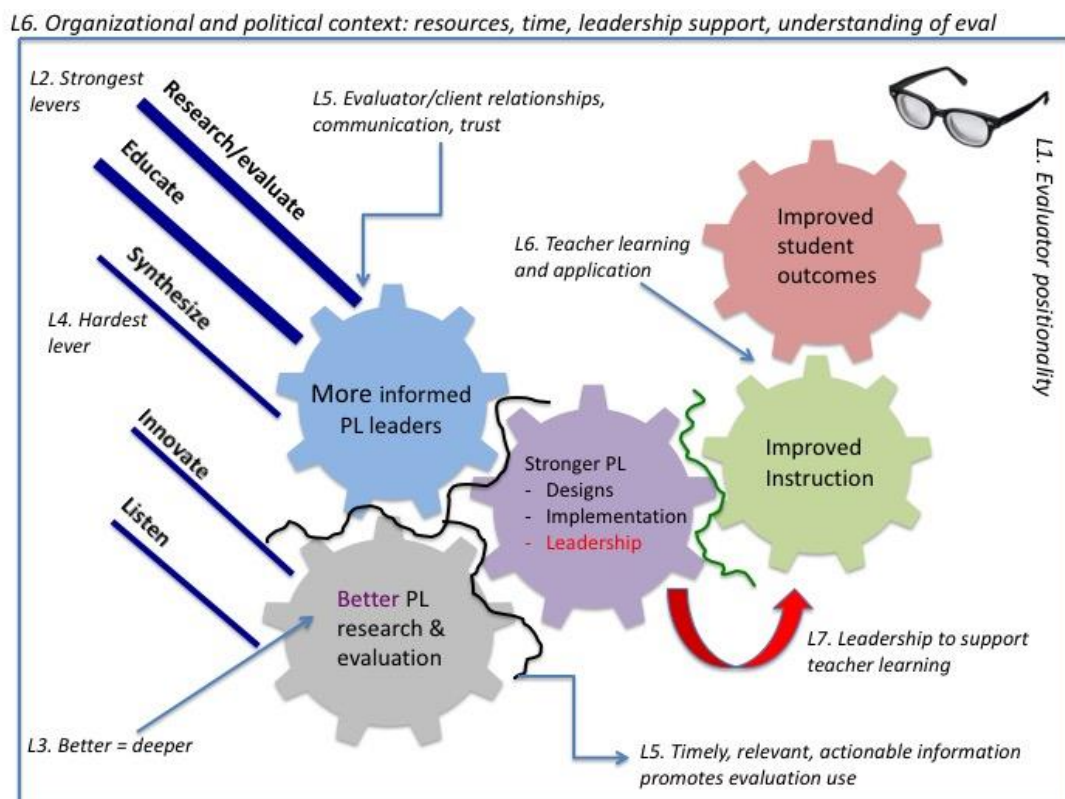


Figure 4 Revised theory of change

As discussed in Chapter 2, my overall theory of change is ambitious. It aims to use strategies originating in one sector (higher education) to influence another sector (Delaware K-12 education). Some might argue that it is not the university's job to improve public schools, although my understanding of UD's mission suggests that it is a key, and mutually beneficial, part of our role. It is challenging to hold ourselves as university staff accountable to improvements in which we participate only indirectly and over which we have no authority. We cannot "make" anyone read or participate in our research, let alone act upon it. But if we stay in touch with and listen to the needs of practitioners, we increase the chances that we can be useful to them. This ELP has strengthened my understanding of how to make evaluation useful.

Chapter 6

REFLECTION ON LEADERSHIP DEVELOPMENT

I have grown in many ways as a leader since I took my first doctoral class in the fall semester of 2013. My EdD classes and projects and my job responsibilities as an educational researcher at DERDC intertwined to create many learning opportunities. Like many EdD students, I also experienced professional transitions over the past four and a half years. Halfway through my program, I decided to step away from my full-time position at UD and instead to combine contract research work for DERDC and DASL with teaching at UD and in the Delaware College Scholars program. In 2017-18, I am additionally serving as a research and professional learning consultant and library staff member at St. Andrew's School. These varied roles have given me an even wider perspective on the issues of professional learning and program evaluation, explored throughout this ELP. My own leadership journey is dynamic. In the rest of this ELP, I chronicle some key moments in my growth as a scholar, problem-solver and partner.

Scholar

As a scholar, I greatly appreciated the flexibility of this program. My advisor and faculty sought to meet my needs and respond to my interests, which somewhat

straddle the boundary between “PhD” and “EdD” thinking. I am drawn to theoretical and conceptual ideas, empirical research, and high levels of academic rigor. At the same time, I value learning with and from practitioners and doing work that is close to the ground and improves educational programs directly. Thus I combined seven EdD core classes with three electives from both the EdD and PhD programs as well as two independent studies and two classes towards the University of Connecticut’s Program Evaluation Certificate. Although idiosyncratic, this course plan helped me master the key academic skills of the EdD program: accessing and/or collecting and analyzing information to guide decision-making and solve educational problems.

I took EDUC 850 as a substitute for EDUC 846. In that class, I honed my use of qualitative research methods, especially observations/field notes and interviews. I applied these skills to my program evaluations at DERDC, which often required me to observe professional learning activities and interview participants. In EDUC 850 I discovered a particular passion for interviewing. I gained further experience in an independent study with Dr. Rosalie Rolon-Dow, in which I conducted interviews with UD undergraduates for a study of racial literacy among pre-service teachers. This project also taught me more about analyzing and interpreting qualitative data, including using Dedoose software. I used all these skills in my study of the TLI, in which I independently conducted and analyzed 27 interviews, a large data set, in a compressed period of time. Although my expertise became primarily qualitative, I drew on quantitative research methods from EDUC 828 and EDUC 827 as well. Some are evident in my evaluation of the Title II program.

Program coursework also helped me access, comprehend, evaluate, and synthesize educational research. I used the template that Dr. Farley-Ripple introduced

to us in EDUC 828 in all my subsequent classes, as well as in my professional work for DERDC. In EDUC 807, our final project was an integrated literature review about a topic of our choice, and I focused on how teachers learn to teach writing. In addition to learning to synthesize research, I also increased background knowledge about professional learning in literacy. Similarly, I geared projects towards my ELP topic of professional learning in EDUC 897, EDUC 818, EDUC 890, EDUC 850, EDUC 891 and in the program evaluation course I took as a substitute for EDUC 863. The first three became artifacts for this portfolio.

Finally, I appreciated the opportunity to think deeply about professional learning and program evaluation. Although these inquiries did not always relate directly to my ELP or my professional projects, they raised new questions and sometimes troubled my understanding in ways that were ultimately productive. For example, in EDUC 852, I learned about a variety of different approaches to research including constructivist and even postmodern program evaluation. These dovetailed with my qualitative interests and, along with context gained through my two classes in the University of Connecticut's graduate certificate, contributed to my knowledge of the field of program evaluation. Similarly, in EDUC 850 I delved into research about professional learning that employed a feminist or critical lens, and grappled with questions like, who decides what teachers should learn? What are the power dynamics in professional development? These contributed to my understanding of TLI, a grassroots initiative with very different structure and dynamics from most of the programs I had previously evaluated. These conceptual shifts caused me to change some of my approaches. For example, instead of simply accepting a program's official

goals and objectives as the only salient ones, I might ask participants what they understood the program to be about.

Problem-Solver

I grew as a problem solver by actually conducting evaluations of professional learning, in the course of which I inevitably faced both practical and conceptual dilemmas. As I discuss in Artifact K (Appendix K), a common challenge was determining how to conduct the highest-quality evaluation given the constraints of time, money, staffing, external requirements etc. Managing an evaluation or research study -- as I did with both of the studies included in this ELP -- is rife with problems to solve. To give two examples illustrated in my artifacts, I tried to increase DERDC's capacity to conduct evaluations by mentoring our graduate research assistants (Artifact H, Appendix H). I also had to pull together disparate professional learning activities into a single evaluation report (Artifact A, Appendix A).

Another problem was working with clients who were skeptical of or resistant to program evaluation and/or who had little background understanding of professional learning. This was more of a political and interpersonal dilemma. Artifact K (Appendix K) explores how I tried to handle it and how I could have done so more effectively. In general, the idea of using multiple perspectives to understand and solve organizational problems -- what in EDUC 891 we called "reframing" (Bolman & Deal, 2013) -- was highly relevant in all of my work. Indeed, that is the very nature of program evaluation, at least as my colleagues and I practiced it. And my ELP as a whole demonstrates my understanding that complex problems, like the general inadequacy of most professional learning programs and evaluations, require multi-faceted solutions.

Partner

Effective program evaluation requires partnering with a wide range of stakeholders: program managers, funders, participants, and more. I get better at this every time I do it. Over the past five years I have enjoyed the benefit of many diverse assignments and client dynamics, some of which are included in this ELP. Artifact K (Appendix K) reflects on myself as a partner in an evaluation context and demonstrates my desire to further improve my skills. Overall my clients viewed me as credible and able to build rapport and learn quickly. They appreciated that I was straightforward and candid and asked them questions to help them think. These are all useful partnership skills.

Most research and evaluation projects also require internal partnership, i.e. working as a team. Although I had some experience in this area before starting the ELP program, I gained more over the past several years. Within DERDC, I partnered effectively with graduate students (as shown in Artifacts/Appendices A and H) and with colleagues, especially our director Dr. Joan Buttram (as shown in Artifacts/Appendices B and C). My collaborations often deepened and my role expanded over time. For example, I first had Dr. Rolon-Dow as a professor in EDUC 850, then asked her to advise me in an independent study, then was invited to join her research team. I first worked with Dr. Chrystalla Mouza as an evaluation client, then was invited to be her preceptor for EDUC 774.

My experience in the EdD program allowed me to expand my professional learning and research networks. As regularly happens in Delaware higher and public education, I often found myself wearing many and overlapping hats. I was in EdD classes with school leaders with whom I had worked at DASL and/or DERDC. I taught some students in EDUC 774 who were in the EdD program with me. I gained

access to some research or evaluation participants because I knew them, or their schools or colleagues, from other situations. I had opportunities to share my work products (for instance, the practice brief) because of prior professional and/or academic networks. I appreciated especially the classmates and colleagues who kept me grounded -- usually, school leaders in K-12 public schools. They were comfortable pressing me for how my ideas could be useful for real educators and the real, rather than the university, world. They told me when my work seemed out of touch or unrealistic. My improvement strategy relies on efforts originating in a university to make sense in, and be useful to practitioners in, a public school context. Partnerships are integral to this model. One of the most important partners a university-based researcher can have is a good friend in a public school who will be very blunt.

I also learned how to build partnerships in less familiar contexts. Because of my prior experiences as a teacher and DASL Associate, I felt most at home in Delaware K-12 schools and with public educator clients. Although EdD classes included strong representation from this group, they also allowed me to meet and learn from higher education administrators and instructors, school board members, educators from other states, and even people in other professions. At the same time, some of my DERDC projects involved working with clients in higher education, often faculty members in the hard sciences. I initially found this intimidating but learned that I could transfer some of my partnership skills, such as being curious about the other person's context and concerns, learning key vocabulary, respecting others' expertise, being straightforward about what you don't know, and listening well.

Finally, as my program comes to an end I am starting to transfer some of my insights into the new context of independent schools. Most recently, I have been

working as a research and professional learning consultant to St. Andrew's School, a boarding high school where I began my career in teaching and where my husband is a teacher and administrator. I used St. Andrew's as a case study for some of my EdD projects (e.g., I conducted an organizational case study of the St. Andrew's Summer Institute for faculty in EDUC 891, and designed a study of student learning outcomes for EDUC 828). Compared to the Delaware public schools, St. Andrew's is a very different context for professional learning. As a highly successful independent school, St. Andrew's is exempt from many of the accountability structures (e.g., local, state or federal oversight; standardized testing; funding or grant requirements) that govern public education or other non-profits. This ironically results in a lack of experience with evaluation or data-based decision-making in general. St. Andrew's also has a traditional culture around professional learning with limited resources devoted to it. Being a fully residential school presents unique challenges for adult learning and collaboration. All of this gives me an opportunity to apply my scholarship, problem-solving and partnership skills in new ways.

Indeed, my EdD experiences equip me with the confidence to enter new educational contexts and address new challenges. My work in the program and in my other UD roles has given me a solid understanding of research and best practices about professional learning, and how to evaluate it. Yet as this ELP demonstrates, my style is not to apply a body of knowledge to a problem in a one-dimensional way. Rather, it is to use what I know, what I can do, and what I can learn in flexible and responsive ways to understand all the aspects of a problem. My roles rarely provided me direct access to taking action to solve problems. Rather, I used relationships and communication skills to inform and influence others.

REFERENCES

- Bennett, T. (2017, September 26). *Professional learning in Delaware*. In Learning Forward webinar, *Speak Up for Title II: Evidence and Evaluation* [webinar]. Retrieved November 19, 2017 from <https://learningforward.org/learning-opportunities/webinars/webinar-archive>.
- Blank, R.K. & de las Alas, N. (2010). Effects of teacher professional development on gains in student achievement: how meta analysis provides scientific evidence useful to education leaders. Society for Research on Educational Effectiveness.
- Bolman, L.G. & Deal, T.E. (2013). *Reframing Organizations: Artistry, Choice & Leadership, 5th Edition*. San Francisco, CA: Jossey-Bass.
- Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educational Researcher*, 33 (8), 3 – 15.
- Calvert, L. (2016). Moving from compliance to agency: What teachers need to make professional learning work. Oxford, OH: Learning Forward and NCTAF.
- Clarke, D. & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18, 947 - 967.
- Coldwell, M. & Simkins, T. (2011). Level models of continuing professional development evaluation: A grounded review and critique. *Professional Development in Education*, 37 (1), 143 – 157.
- Combs, E. & Silverman, S. (2016). Bridging the gap: Paving the pathway from current practice to exemplary professional learning. Frontline Research & Learning Institute. Retrieved October 21, 2017 from https://www.frontlineeducation.com/FrontlineEducation/media/images/Resources/ESSA_Bridging_the_Gap.pdf
- Darling-Hammond, L., Hyler, M.E. & Gardner M. *Effective Teacher Professional Development*. Palo Alto, CA: Learning Policy Institute. Retrieved October 29, 2017 from: https://learningpolicyinstitute.org/sites/default/files/product-files/Effective_Teacher_Professional_Development_REPORT.pdf
- Delaware Department of Education (2016a). Reimagining professional learning innovation grant application. Retrieved October 23, 2016 from: <http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/95/CIPD%20PL%20Grant%20FY16%20v4.pdf>

- Delaware Department of Education (2017a). State of Delaware school profile. Retrieved October 20, 2017 from: <http://profiles.doe.k12.de.us/SchoolProfiles/State/Account.aspx>
- Delaware Department of Education (2017b). State assessment results show more students proficient in math and English. [Press release].
- Delaware Department of Education (January 3, 2014). Highly qualified teachers and improving teacher quality state grant program [Request for proposal].
- Desimone, L. (2009). Improving impact studies of teachers' professional development: Towards better conceptualizations and measures. *Educational Researcher*, 38 (3), 181 – 199.
- Doino-Ingersoll, J.A., Haley, J., Dowell, K., & Chambliss, J. (2005). Client feedback form. Retrieved May 26, 2017 from: <https://tinyurl.com/yc5jq6fc>
- Evans, L. (2014). Leadership for professional development and learning: Enhancing our understanding of how teachers develop. *Cambridge Journal of Education*, 44 (2), 179 – 198.
- Furco, A. (2013). Legitimizing community engagement with K-12 schools. *Peabody Journal of Education*, 88 (5), 622 – 636.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F. & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38 (4), 915 – 945.
- Bill & Melinda Gates Foundation (2014). Teachers know best: Teachers' views on professional development. Retrieved November 22, 2016 from: <http://k12education.gatesfoundation.org/wp-content/uploads/2015/04/Gates-PDMarketResearch-Dec5.pdf>
- Guskey, T.R. (2002). Professional development and teacher change. *Teachers and Teaching*, 8 (3), 381-391.
- Guskey, T.R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership*, 59 (6), 45 – 51.
- Guskey, T. R. (2000). *Evaluating professional development*. Newbury Park, CA: Corwin Press.

- Kennedy, A. (2014). Understanding continuing professional development: The need for theory to impact on policy and practice. *Professional Development in Education*, 40 (5) 688-697.
- Kennedy, M.M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86 (4), 945 – 980.
- Kose, B.W. (2009). The principal's role in professional development for social justice. *Urban Education*, 44 (6), 628-663.
- Layton, L. (2015, August 4). Study: Billions of dollars in annual teacher training is largely a waste. The Washington Post. Retrieved November 17, 2016 from: https://www.washingtonpost.com/local/education/study-billionsof-dollars-in-annual-teacher-training-is-largely-awaste/2015/08/03/c4e1f322-39ff-11e5-9c2ded991d848c48_story.html
- Learning Forward (2015). Definition of professional development. Retrieved November 22, 2016 from: <https://learningforward.org/who-we-are/professional-learning-definition>
- Learning Forward (2011). Standards for professional development. Retrieved April 21, 2014 from <http://learningforward.org/standards>
- National Research Council (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- New Teacher Center (2013a). TELL Delaware results. Retrieved October 23, 2016 from: <https://telldelaware.org/results/report/80/1346#PD>
- New Teacher Center (2013b). How different educators perceive teaching conditions. Retrieved October 23, 2016 from: https://telldelaware.org/uploads/File/DE13_brief_comparisons.pdf
- Opfer, V.D. & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, 81 (3) 376 - 407.
- Penuel, W.R., Fishman, B.J., Yamaguchi, R. & Gallagher, L.P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, 44 (4), 921 – 958.

- Robertson-Kraft, C. , Hoe, N., Sangenito, S. & Williams, E. (2017). Excellent educators report. Philadelphia, PA: Operation Public Education. Retrieved October 20 2017 from https://www.doe.k12.de.us/cms/lib/DE01922744/Centricity/domain/37/tleu%20presentations%20and%20reports/Delaware%20Excellent%20Educators_Final%20Report.pdf
- Rogers, P. (2014, May 9). Ways of framing the difference between research and evaluation. (Web log post). Better Evaluation. Retrieved February 12, 2018 from: http://www.betterevaluation.org/en/blog/framing_the_difference_between_research_and_evaluation
- Schools That Lead (2017, June 12). Measuring impact. [Online subscription newsletter article].
- Strategic Data Project (2015). SDP Educator Diagnostic for the Delaware Department of Education. Retrieved January 29, 2018 from: <https://www.doe.k12.de.us/cms/lib/DE01922744/Centricity/domain/37/tleu%20presentations%20and%20reports/SDP%20Educator%20Diagnostic%20Delaware.pdf>
- Stufflebeam, D.L. (1999). Program evaluations meta-evaluation checklist. The Evaluation Center, Western Michigan University. Retrieved March 16, 2017 from: https://www.wmich.edu/sites/default/files/attachments/u350/2014/program_metaeval_short.pdf
- Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D.B. & Balcazar, F.E. (2013). Understanding and measuring evaluation capacity: A model and instrument validation study. *American Journal of Evaluation*, 34 (2), 190 – 206.
- Timperley, H. (2011). *Realizing the power of professional learning*. Maidenhead, England: Open University Press.
- TNTP (2015). The mirage: Confronting the hard truth about our quest for teacher development. Retrieved November 22, 2016 from: <http://tntp.org/publications/view/the-mirage-confronting-the-truth-about-our-quest-for-teacher-development>
- Walsh, M.E. & Backe, S. (2013). School-university partnerships: Reflections and opportunities. *Peabody Journal of Education*, 88 (5), 594 – 607.

- University of Delaware (2014). 2015 Documentation reporting form: Carnegie Community Engagement Classification (First Time Applicants). Retrieved January 6, 2017 from: <https://sites.udel.edu/engage/files/2016/08/University-of-Delaware-first-time-app-11elobr.pdf>
- University of Delaware (2015). Delaware will shine: A blueprint for a pre-eminent learner-centered research university. Retrieved January 6, 2017 from: <http://www1.udel.edu/research/delawarewillshine/DelawareWillShine-FINAL.pdf>
- Weiss, C. (1998). *Evaluation* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Wiggins, G. & McTighe, J. (2006). *Understanding by design* (2nd Ed.) Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- Yarbrough, D.B., Shulha, L.M., Hopson, R.K., & Caruthers, F.A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users*. Thousand Oaks, CA: SAGE.
- Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>

Appendix A

ARTIFACT A: EVALUATION OF THE 2015 ACTIVITIES OF THE DELAWARE TITLE II GRANT

Evaluation of the 2015 Activities of the Delaware Title II Grant,

**Professional Development to Support Quality Teaching and
Student Learning in High-Need Schools**

**Hilary Mead
Amanda Czik
Jordana Woodford
Rachel Salinger**

**Delaware Education Research & Development Center
September 2015**



Delaware Education Research & Development Center

University of Delaware

Newark, DE 19716

PUBLICATION T2015.11

Executive Summary

The Delaware Title II grant aims to build the pedagogical and content knowledge of teachers working in high-need schools in Delaware and thus to support teacher effectiveness and, ultimately, student achievement. In 2014-15, the project was a partnership between the University of Delaware (UD), three school districts, and two charter schools. In general, the project aimed to help educators and students make the shift to new, more rigorous academic standards and to provide high-quality professional development (PD) aligned to the new standards and assessments.

This year's grant included seven distinct PD efforts, spanning multiple content areas and demonstrating a variety of designs:

- Individual coaching in guided reading strategies
- Curriculum development of writing modules
- Summer institute on non-fiction writing
- Social studies webinar series
- *Responsive Classroom* PD and coaching
- Educator Design Workshop (hands-on STEM)
- Alternative routes to certification for secondary special education/world language teachers

The Delaware Education Research & Development Center (DERDC) conducted an evaluation of the Title II activities to answer the following questions:

1. Did the project reach its target audience?
2. Was the PD provided high quality? Were participants satisfied?
3. Did teachers develop the target content and pedagogical content knowledge?
4. Does teachers' instruction demonstrate the target skills/knowledge?
5. Does student achievement in the target skills/knowledge improve?

Data to answer these questions came from demographic surveys, evaluator observations, participant surveys, and review of PD materials and participant work products. For activities that occurred over time and/or in particular schools, participant and administrator interviews were also conducted. Where possible, the evaluation also included secondary data related to instructional change and student achievement.

Overall, the evaluation found that the activities reached their target audience and far exceeded the target number of participants. Participants were on the whole satisfied, with 82.8% overall agreeing or strongly agreeing that the PD was a good use of their time. Satisfaction varied within and across activities, particularly *Responsive Classroom*, but this was a fundamentally more ambitious effort than the others because it involved school-wide participation. It also accounted for over 75% of all Title II participants.

The PD activities focused on content and pedagogical content knowledge, and aligned well to the CCSS, NGSS, and Smarter Balanced Assessment. They also consistently emphasized collaboration. Looking across all seven activities, we see several research-supported PD designs, including one-on-one coaching, hands-on active learning, and involvement of teams and administrators. However, the activities varied widely in design.

Self-reports of participant development were generally positive. The PD activities modeled a range of instructional techniques that participants could observe, try out, and bring back to their classrooms. Most but not all activities provided time and support for participants to plan for application, but the scope varied considerably. In some activities, participants submitted plans that could be evaluated; these tended to be inconsistent.

Evaluation of instructional change and student achievement was possible for four of the seven activities although only three are reported due to the small size of DARE. The results generally indicate efforts at implementation of new practices but mixed results. Our data affirm that instructional change does not happen overnight and requires ongoing support from administrators and professional developers. Student achievement targets were mostly met in reading (East Side Literacy Coaching) and writing (writing module development). School climate data (*Responsive Classroom*) were mixed. The report ends with suggestions for strengthening future Title II grants.

Table of Contents

Introduction	68
Evaluation Design	69
Evaluation Procedures	69
Findings	70
Evaluation Question 1: Target Audience	70
Evaluation Question 2-5: Participant Satisfaction, Participant Learning, Instructional Change, and Student Achievement	72
Literacy Coaching at East Side Charter School.....	73
Writing Module Development.....	80
Summer Writing Institute	89
Social Studies Webinar Series.....	96
<i>Responsive Classroom</i>	102
Educator Design Workshop	113
DARE	120
Discussion.....	124
Recommendations	126
References.....	126
Appendix 1: Standards.....	127
Appendix 2: Participant Demographics	130
Appendix 3: Sample Participant Survey	132
Appendix 4: Participant Interview Protocol	134
Appendix 5: Administrator Interview Protocol	135
Appendix 6: Writing Module Checklist	136
Appendix 7: Lesson Plan Rubric.....	137

Introduction

The Delaware Title II grant aims to build the pedagogical and content knowledge of teachers working in high-need schools in Delaware and thus to support teacher effectiveness and, ultimately, student achievement. In 2014-15, the project was a partnership between the University of Delaware (UD), three school districts, and two charter schools. In general, the project aimed to help educators and students make the shift to new, more rigorous academic standards and to provide high-quality professional development (PD) aligned to the Common Core State Standards (CCSS) or the Next Generation Science Standards (NGSS).

The project included seven distinct activities, each with a defined goal. These activities spanned various content areas and used different PD models, as summarized in Table A1.

Table A1
Overview of Title II Program Activities in 2014-15

Goal	Activity & Professional Development Model	Timeframe
Improve teachers' understanding and implementation of guided reading	Literacy coaching at East Side Charter School One-on-one modeling, observation and coaching	October 2014 through May 2015
Improve teachers' content and pedagogical knowledge of writing	Writing curriculum development Teams of teachers in grades 3, 5 and 9 developing integrated writing modules, in collaboration with DDoE	December 2014 through May 2015
	Summer institute on non-fiction writing	July 2015
Improve teachers' content and pedagogical knowledge of civics and history	Series of online PD webinars	December 2014 through May 2015
Improve teachers' understanding and implementation of Responsive Classroom techniques	Partnerships with elementary or K-8 schools, mixture of PD workshops and coaching	October 2014 through June 2015
Improve teachers' content and pedagogical knowledge in STEM	Series of design workshops including three Saturdays and a three-day summer workshop	March 2015 through July 2015
Recruit, select, train, develop and retain highly qualified teacher candidates in secondary education	Alternative route to teacher certification program for candidates in world languages or special education	October 2014 through July 2015

Across all these activities, the project aimed to provide high quality PD to support effective teaching, as set forth in state and national standards (see Appendix 1 for standards).

Evaluation Design

The Delaware Education Research & Development Center (DERDC) conducted an evaluation of Title II activities to determine the extent to which the project achieved its goals and exemplified high-quality PD. The proposed evaluation was submitted to and granted exempt status from UD's Institutional Review Board. The design followed Guskey's (2012) model for PD evaluation focused on the following evaluation questions:

1. Did the project reach its target audience?
2. Was the PD provided high quality? Were participants satisfied?
3. Did teachers develop the target content knowledge and pedagogical content knowledge?
4. Does teachers' instruction demonstrate the target skills/knowledge?
5. Does student achievement in the target skills/knowledge improve?

Data to address these questions came from a variety of sources as described in the next section.

The seven PD activities were similar in some ways and very different in others. In particular, they varied in scope, both the number of participants and the span of time over which activities occurred. In conducting this evaluation, we tried to balance consistency and particularity. We asked the same evaluation questions for each activity and went into more depth where data were available to support our inquiry.

Evaluation Procedures

Demographic information on Title II participants was gathered through a Delaware Department of Education (DDoE) form, which was administered at the end of each activity. The form included questions about the participant's professional role, district/charter affiliation, licensure status, educational level, years of experience, clock hours, reason for participating in the professional development, race/ethnicity, and gender. The form addressed Evaluation Question 1. It also collected data on the number of students reached by each participant. Demographic data were analyzed quantitatively using SPSS. Complete results are provided in Appendix 2.

An evaluator attended multiple sessions of each PD activity and conducted observations using a protocol based on PD standards. Copies of agendas and other materials were obtained and reviewed for alignment to these standards, as well.

A participant survey was also administered on the last day of each institute, but data from the two surveys could not be matched. The survey instrument was similar for each activity, with a few items

tailored to its specific content.¹ See Appendix 3 for a sample. It measured participant satisfaction and ratings of PD quality and effectiveness to answer Evaluation Question 2. There were on average 19 Likert scale and five open-ended items. These questions asked participants what they found most useful and what they would change; how the PD strategies aligned with their practice; and what challenges they anticipated in using what they learned. Frequencies, means, and standard deviations were calculated for the ratings items. Open-ended responses were analyzed qualitatively, coding for themes. This report summarizes qualitative data with exemplars. The complete text of the open-ended responses has been shared with project staff. Where an activity occurred over time, results from each set of surveys were shared to facilitate formative improvement (e.g., we shared *Responsive Classroom* survey results from each school as they became available).

For all activities, the participant survey also included questions about participant development (Evaluation Question 3). Other measures of teacher growth in skills and knowledge (Evaluation Question 3) and instructional change (Evaluation Question 4) varied from activity to activity. These included review of participant work products; pre/post assessments; and participant and administrator interviews. Finally, the evaluation team collected available data on student learning (Evaluation Question 5) including writing scores, school climate, and student retention data. In general, activities that were ongoing over time and focused on a particular school were able to be evaluated more intensively through these methods. For activities that occurred during summer 2015, insufficient time elapsed to measure the impact on instructional practice and student learning. Further details on data collection procedures and instruments specific to each activity are provided in the sections below.

Findings

The findings section is set up as follows. First, information is presented about the participants in the Title II activities as a whole, including enrollment numbers and demographics. This addresses Evaluation Question 1. We also consider the total number of *students* that these participants reached, as one measure of the grant's impact on student learning. In subsequent sections, each activity is briefly introduced, then findings are presented related to PD design (Evaluation Question 2), participant satisfaction (Evaluation Question 2), participant learning (Evaluation Question 3) and instructional change (Evaluation Question 4). Where the activity occurred during the 2014-15 school year and student level data are available, these are also included (Evaluation Question 5). Due to differences in scope, the sections are not all the same length. The Discussion section synthesizes the findings across all activities.

Evaluation Question 1: Did the project reach its target audience?

The target audience for Title II activities was teachers in high-need schools and a goal was set to reach 200 of them. Enrollment numbers and responses to the DDoE demographic forms show that the project far exceeded this goal. As shown in Table 2, according to project records the enrollment across all activities totaled 376.

¹ In keeping with the activity format, the social studies survey was administered online and used slightly different items.

Table A2
Title II Institute Enrollments

Activity	Number Enrolled	Percentage Enrolled
Literacy Coaching at East Side Charter School	10	2.6
Informative Writing Module Development ^a	11	5.9
Summer Non-Fiction Writing Institute	17	4.4
<i>Responsive Classroom</i> PD/Coaching ^b	292	75.3
Social Studies Webinars ^c	29	7.5
Educator Design Workshop	11	2.8
DARE	6	1.5
Total	376	100.0

^a This activity also included 12 participants who were funded through a different grant. Only the Title II-supported participants were included in our demographics.

^b Does not include participants from Oberle, Shortlidge or Warner Elementary Schools as these did not complete RC activities fully and no participant or demographic surveys were collected in these schools.

^c Only includes the 29 participants who attended at least one webinar live. A total of 87 educators registered for a webinar but it is not known how many accessed the recording after the webinar. Thus we are using a conservative estimate for this participation number.

Demographic data for participants are in Appendix 2. These are based on the 336 demographic forms received. There are two reasons why the number enrolled (376) exceeds the number of demographic forms (336). Numerous participants attended more than one Title II activity but were only asked to complete one demographic form each. For instance, at least 20 participants attended *Responsive Classroom* PD through their schools during 2014-15 and then attended a summer institute. Ten educators from East Side Charter School participated both in literacy coaching and *Responsive Classroom*. Second, demographic forms were not obtained from all participants. If a candidate was missing on the last day of the activity, s/he did not complete this form. In addition, demographic forms were sent electronically to participants in the social studies webinars, but not all completed it.

Almost all participants were public school educators (95.2%). Most were teachers (86.8%) with the rest administrators (4.8%), paraprofessionals (1.8%) and "other" (6.6%). The majority (62.2%) had a master's degree. Their years of teaching experience ranged from 0 (graduates entering their first jobs in 2015-16) to over 30, but the largest group (38.4%) had less than five years of experience. The majority (79.6%) was from elementary schools, with 10.5% from middle schools, 5.2% from high schools, and 4.6% working with multiple ages. Over half the teachers (52.3%) reported serving 30 students or fewer. On the other end of the distribution, 19.9% served 101 students or more. In terms of personal demographics, the majority of participants were white non-Hispanic (79.0%) and female (88.1%).

Participants were asked to identify their district/charter school affiliations (see Table 3, below).

Table A3
District or Charter Affiliations of Title II Participants

District or Charter	Number	Percent
Colonial School District	143	43.5
East Side Charter School	51	15.5
Las Américas ASPIRA Academy	43	13.1
Red Clay Consolidated School District	29	8.8
Christina School District	17	5.2
Other district	31	9.4
Other charter school or charter school, not specified	15	4.6

N=329. Seven participants did not respond to this item. Percents do not add to 100 due to rounding.

The three districts and two charters partnering on this grant all participated in activities but to varying extents. Over 40% of the Title II participants this year came from the Colonial School District. This was because three elementary schools in this district required school wide participation in *Responsive Classroom*; some Colonial educators attended other Title II activities, also. ASPIRA and East Side also had school wide participation in *Responsive Classroom*. There was less participation from the Christina and Red Clay School Districts. In addition, 14% of participants were employed by a non-partnership district or charter school (or indicated they worked for a charter but did not name it).

Altogether, numeric responses to the demographic survey suggest that the institutes indirectly reached over 53,503 students. This number is an estimate using the *lowest* range from all participant self-reports of how many students they reached and may also reflect that some students from the same districts may be accounted for more than once. This is the total potential impact on students. However, we recognize that some of these students interacted more closely with institute participants than others. An elementary school teacher would “reach” a smaller number of students than, for example, a district instructional coach, but s/he would spend more time with each of those students. Excluding responses from administrators, the total number of students reached is 30,749.

Evaluation Questions 2 – 5:

Was the PD provided high quality? Were participants satisfied?
Did teachers develop target content and pedagogical content knowledge?
Does teachers’ instruction demonstrate the target skills/knowledge?
Does student achievement in the target skills/knowledge improve?

The next six sections answer these evaluation questions for each individual activity. These are presented in the order they were presented in the proposal and in Table A1 on page 2, above. In each case, we briefly introduce the activity, the evaluation methods, and the participants. Then we separate findings into participant satisfaction and PD design, participant learning and, where data are available, instructional change and student achievement.

Literacy Coaching at East Side Charter School

Introduction. East Side Charter School, a high-needs school in Wilmington, partnered with UD to support reading instruction in the primary grades. A coach from UD provided one-on-one modeling, observation, and feedback to all teachers in kindergarten, first and second grade (total: 10). She focused on guided reading, an instructional approach that uses small groups and teaches reading strategies in the context of reading authentic books and stories.

There were two rounds of coaching, the first in October/November 2014 and the second in March/April 2015. In the first round, the coach modeled guided reading in each classroom, and then observed the teacher and offered coaching and feedback as she tried the strategies herself. The second round was intended as a progress check.

Data for this activity were collected as follows. Teachers completed a survey in late November, after the first round of coaching, to permit mid-course adjustments. After the second round of coaching, all teachers participated in one-on-one interviews with a member of the evaluation team as did a school administrator. Finally, the school submitted achievement data related to student reading growth and promotion to the next grade.

Participant satisfaction and professional development design. After the first round of coaching, participants rated the PD experience (Table A4).

Table A4
Participant Ratings of Literacy Coaching at East Side Charter School Design

The literacy coaching:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
1. Used materials that are relevant to my teaching situation.	6 (60.0%)	4 (40.0%)	0 (0.0%)	0 (0.0%)	3.60 (0.52)
2. Had a knowledgeable coach	7 (70.0%)	3 (30.0%)	0 (0.0%)	0 (0.0%)	3.70 (0.48)
3. Gave me sufficient opportunities to ask questions.	8 (80.0%)	2 (20.0%)	0 (0.0%)	0 (0.0%)	3.80 (0.42)
4. Had sufficient time allocated to it. *	6 (66.7%)	2 (22.2%)	1 (11.1%)	0 (0.0%)	3.56 (0.73)
5. Modeled effective instruction.	6 (60.0%)	3 (30.0%)	1 (10.0%)	0 (0.0%)	3.50 (0.71)
6. Addressed a topic that was new to me.	4 (40.0%)	5 (50.0%)	1 (10.0%)	0 (0.0%)	3.30 (0.67)
7. Responded to my needs as a teacher.	6 (60.0%)	4 (40.0%)	0 (0.0%)	0 (0.0%)	3.60 (0.52)
8. Responded to the needs of my students.	7 (70.0%)	2 (20.0%)	1 (10.0%)	0 (0.0%)	3.60 (0.70)
9. Fit the way I learn best.	6 (60.0%)	4 (40.0%)	0 (0.0%)	0 (0.0%)	3.60 (0.52)

The literacy coaching:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
10. Gave me enough opportunities to practice.	6 (60.0%)	3 (30.0%)	1 (10.0%)	0 (0.0%)	3.50 (0.71)
11. Was worth my time.	8 (80.0%)	2 (20.0%)	0 (0.0%)	0 (0.0%)	3.80 (0.42)

N=10.

*One respondent skipped this question so N=9.

The majority of these ratings were positive. Two individuals were responsible for the negative ratings. Teachers felt particularly strongly that they had enough opportunities to ask questions (item 3) and that the coaching was worth their time (item 11).

Open-ended survey items asked participants to describe their favorite aspects of the coaching and to provide suggestions for improvement (Tables A5 and A6).

Table A5

Participants' Favorite Aspects of Literacy Coaching at East Side Charter School

Response (Category)	Number	Percentage
Support, modeling and opportunities for practice	6	75.0
Varied strategies and new ideas	4	50.0
Realistic focus on implementation	3	37.5
Strategies that will work for students	3	37.5
Specific information (learning how to choose books)	1	12.5

N=8. Two participants skipped this question. Some respondents included more than one response.

Table A6

Participants' Suggestions for Improvement for Literacy Coaching at East Side Charter School

Response (Category)	Number	Percentage
More time for the coaching process	4	50.0
Nothing/no suggestions	3	37.5
More encouragement/positive feedback	2	25.0

N=8. Two participants skipped this question.

Most of the positive comments affirmed the one-on-one coaching model and the value of modeling. One teacher appreciated "getting to see an expert work w/my kids. It gave me so many ideas." Another stated she "benefitted from the measurable and bite-sized action steps. [They] were realistic and I could implement them in my classroom the very next day." In terms of suggestions, most teachers only wanted more time or opportunities for coaching although two suggested the need for more positive encouragement: "emphasize praising."

Later in the year, we obtained feedback on the same questions through the teacher interviews. We heard again that teachers appreciated the concrete, specific, and personalized PD model:

I felt like her coming into each individual room made it possible for her to help you make sure it was specific to your room.

Not just being told what to do but being shown is really helpful for people who have done it and for people who never have. And I think it's really helpful in that it's not a video or fishbowl...it's your kids. So if someone can work with them in this way who doesn't know them, then what are you going to be able to do?

All comments about the coaching style were positive. Words like “open”, “honest” and “supportive” came up often. Teachers generally felt they had a strong relationship with the coach and could contact her with questions or for resources. One called her, “by far the best coach I have had all year.” The coach gained respect because of her ability to quickly grasp the school and student context; one teacher noted that she learned students’ names, another that she was not phased by discipline issues.

We also sought feedback for improving the coaching model. Most interviewees suggested simply that there should be more coaching visits or that the second visit should happen sooner (“versus the end of the year, after spring break, when it’s like, ‘sorry, just 20 days left’”). Some teachers still appeared uncertain about getting all the components of guided reading right and wanted more reinforcement. For example, one requested a “scope and sequence” for how students learn to read. Another suggested that it would be more beneficial to have an intensive PD experience before the school year, with follow-up coaching afterwards. A third suggested that paraprofessionals also receive PD or coaching so they could better support guided reading.

Participant learning. Here we present results from the mid-year survey and later we turn to the interview data. Participants rated a series of statements about workshop outcomes; see Table 7, below.

Table A7

Participants’ Ratings of the Literacy Coaching at East Side Charter School Learning Outcomes

After participating in the literacy coaching:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
12. I learned how to select appropriate books for my different groups of readers.	2 (20.0%)	8 (80.0%)	0 (0.0%)	0 (0.0%)	3.20 (0.42)
13. I learned how to embed skill instruction into reading stories.	3 (30.0%)	7 (70.0%)	0 (0.0%)	0 (0.0%)	3.30 (0.42)
14. I have seen examples of strategies that will work with my students.	5 (50.0%)	5 (50.0%)	0 (0.0%)	0 (0.0%)	3.50 (0.53)

After participating in the literacy coaching:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
15. I know where to access resources/ materials to support my reading instruction.	2 (20.0%)	8 (80.0%)	0 (0.0%)	0 (0.0%)	3.20 (0.42)
16. Lack of resources at my school will make the strategies we covered difficult to implement.	1 (10.0%)	2 (20.0%)	7 (70.0%)	0 (0.0%)	2.40 (0.70)
17. I feel confident in my ability to implement strategies from the coaching with my students.	4 (40.0%)	6 (60.0%)	0 (0.0%)	0 (0.0%)	3.40 (0.52)
18. I still have many questions about guided reading.	1 (10.0%)	5 (50.0%)	4 (40.0%)	0 (0.0%)	2.70 (0.67)
19. I have discussed the reading strategies with colleagues in my school.	2 (20.0%)	8 (80.0%)	0 (0.0%)	0 (0.0%)	3.20 (0.42)

N=10.

As shown in these ratings, most educators were reporting growth in their understanding of guided reading midway through the year. Ratings of the relevance and usefulness of the reading strategies (item 14) were especially high. At the same time, these ratings indicate potential challenges. Six respondents agreed they “still ha[d] many questions about guided reading” (item 18). The second round of coaching was intended to address these lingering questions. Three anticipated that insufficient school resources might be an issue (item 16). Open-ended responses echo that books, time, and professional development are necessary for implementation (Table A8). Other themes included the challenge of “choosing the best strategies” or “continuing, implementing with fidelity.”

Table A8

Participants’ Anticipated Challenges for Implementing Literacy Coaching at East Side Charter School

Response (Category)	Number	Percentage
Lack of materials, resources or support	3	37.5
Need to choose appropriate strategies	2	25.0
Fidelity of implementation	2	25.0
None come to mind at the moment	2	25.0

N=8. Two participants skipped this question.

Finally, participants were asked how students would respond, or had responded, to guided reading (Table 9).

Table A9

Participants' Expectations for Student Response to Literacy Strategies Demonstrated through Coaching

Response (Category)	Number	Percentage
Students responded well (general)	6	60.0
Students were engaged/excited	4	40.0
Students gained self-efficacy	3	30.0
Students responded well to specific strategy	1	10.0

N=10. Several respondents included more than one response.

Many teachers made general positive statements about how students enjoyed reading, while others described increased engagement and self-efficacy: *"Very well! They are doing a lot more self-correcting now."* More information about students' learning gains is described in the section below.

Instructional change. Data related to instructional change come from the teacher/administrator interviews. The teachers participating in coaching had a range of experience with guided reading. A few had prior coaching or PD and already considered themselves proficient. For others, it was brand new. Most commonly, teachers stated some variation of this comment: *"I was doing guided reading but it really didn't feel like it. Compared to this year, I don't think it really was guided reading."* Coaching helped teachers to deepen, structure, and improve the small group reading instruction they were providing. Teachers across the experience spectrum described the coach meeting them where they were. For example, one described discussing how to play a leadership role with guided reading. Another described the coach as *"not a threat"* even though she was demonstrating new strategies that conflicted with her prior practices.

All interviewees stated that they had learned new methods of reading instruction and were using them in their classrooms. However, there was a range of specificity. Some spoke in general terms: *"she helped me get better at the reading aspect"* or *"small grouping is a lot better than the large grouping."* Others gave detailed examples of strategies they had learned or tried. Most commonly, guided reading coaching helped teachers identify and address student reading levels much more specifically. Many classes had wide ability ranges and several teachers described the coach helping them challenge more proficient readers and/or offer targeted support to struggling students. Teachers described how the coach provided actionable suggestions and tangible resources, such as windows students could use to focus their attention on a single word. Two teachers described putting suggestions into practice on the very same day they had been coached. In general there was consensus that the coach was "crystal clear" about the components of guided reading. While there appear to be widespread attempts to implement guided reading, there is also some evidence that it was inconsistent.

Some teachers have really taken off with it and done amazing and just like anything some teachers aren't implementing it as regularly as they should. And some are taking bits of it but not doing it fully.

Interviewees were asked what organizational conditions facilitated or impeded implementation. The PD model itself was seen as a facilitator. Teachers described other supports, such as tangible instructional resources (e.g., reading books, magnetic letters) and designated time during the school day for guided reading. Teachers also generally found administrators and instructional coaches supportive of guided reading, and some had small class sizes, which were also conducive. The largest challenge was the school's lack of a reading program. This meant that teachers had to invest considerable time and energy into identifying and organizing resources. Some teachers appreciated the flexibility, but others stated simply, *"I wish we could get a program and stick to a program."* Another barrier was classroom management, which sometimes meant that teachers or paras got pulled into handling discipline and were not available to lead small groups.

Finally, some suggestions related to administrative oversight or alignment. This year, several sources stated that there was open communication between the coach and the administrator, who sought to build his own understanding of guided reading. However it was suggested that more could be done to codify such communication:

[UD staff] never say, 'well, on your side, what are you going to do?' but it might be worth it for them to push the school...

So if they were to spell out some agreement or commitment, what should they ask for?

That they will regularly check in on teachers to make sure they are doing [guided reading] at all, that they offer it as part of their coaching as a school and have conversations with UD around alignment.

A few interviews also indicated the need for administrators to monitor guided reading implementation more closely and consistently.

Student achievement. The interviews also explored the impact of guided reading on students. There was consensus that students were more engaged and excited about reading now. Some teachers joked that students did not want to stop reading: *"Sometimes I'm like, 'All right, I'm talking. Put your books away.'"* Others talked about increased student confidence: *"They are like, 'oh my gosh, there are so many words on this page and I'm reading it!'"* Some teachers reported continued struggles with some students (*"my lowest group is my lowest group because things continue not to work"*) but overall there was a sense that student skill growth accelerated in 2014-15. Some seemed almost amazed: *"I could see kids learning to read in a way that just wasn't happening our school before that."* Others focused on impacts on teacher confidence and skill in the complicated process of teaching a student to read.

I don't care if you were trained in elementary education, you weren't trained in how to teach a kid to read. And so just seeing some teachers really know what to do to get a kid to read and not be overwhelmed by the task, especially kids who are already really behind and knowing what's at stake. I think that's the biggest positive to me.

The project established four specific targets for student learning. These are listed in Table A10, along with the relevant data from 2014-15, all of which were collected, analyzed and reported to the evaluators by school personnel.

Table A10

Student Learning Targets and Performance for Grade K-2 Reading at East Side Charter School

Targets	Performance in 2014-15
At the end of kindergarten, at least 25% of students will read on a first grade entry level according to Action 100 scores.	24% of kindergarten students read on a first grade entry level as of June 2015.
At the end of first grade, at least 25% of students will read on a second grade entry level according to Action 100 scores.	25% of first grade students read on a second grade entry level as of June 2015.
The number of students retained in kindergarten due to literacy and reading issues will decrease to 6% or less.	Three students (4.9%) were retained in kindergarten between the 2014-15 and 2015-16 school years.
On the MAP test, all grades K, first and second will average at least 0.75 of a year's growth from fall to spring.	Among first graders (n= 41), the average end-of-year reading MAP growth obtained was 0.88 years. The range was -0.1 years to 2.4 years. MAP data for second graders was not provided in a way that permitted these analyses. Kindergarten students did not take the MAP in 2014-15.

Note: there was also a goal related to the Smarter Balanced Assessment but data were not available in time and the assessment was not administered in grades K-2.

As described in Table A10, the school met most of the targets for which data were available. The rate of students repeating kindergarten for literacy reasons decreased. Kindergarten students approached but did not quite meet the Action 100 reading level target; the difference appears to be just one (of 61) students. In first grade, students met Action 100 target as well as the MAP growth target.

In summary, this activity was highly personalized, offering intensive and ongoing support to a small number of teachers. Satisfaction was pervasively high, and all participants described learning and attempting to use new skills. While the coaching may have moved all participants forward, implementation and monitoring of guided reading still needed more consistency. Data on student outcomes, although incomplete, are promising. The school met three targets for reading growth and approached the fourth.

Writing Module Development

Introduction. The University of Delaware (UD) partnered with the Delaware Department of Education (DDoE) on a curriculum development activity. Teams of regular and special education teachers and librarians collaborated to develop integrated, scaffolded instructional modules on non-fiction writing and research. The purpose was to develop modules that would give all students, with and without disabilities, access to rigorous, CCSS-aligned instruction, thus improving their writing skills. DDoE and UD staff members together facilitated this activity.

The activity included three evening group PD sessions during the winter and early spring. These focused on understanding and planning backwards from the CCSS-ELA standards and learning progressions; analyzing text complexity; developing performance tasks and formative assessments; and using Universal Design for Learning. In between sessions, teams worked to develop their modules, and received feedback and support from the project facilitators. At least one teacher from each team piloted at least one lesson from the module in the spring. Finally, teams presented their completed modules to the Delaware Literacy Coalition in May 2015. After completing these requirements, each participant received a stipend.²

The five teams were school- and grade-level based. There were two in third grade (Bunker Hill and Clayton Elementaries), two in fifth grade (Lulu Ross and Leasure Elementaries) and one in ninth (Smyrna High). Teams ranged in size from three participants (Clayton) to six participants (Smyrna and Lulu Ross). Across the five teams, there were a total of 23 participants. All but one (n=22) attended the final presentations and completed surveys about their experience.

In addition to these satisfaction surveys, teachers also completed surveys related to content and pedagogical knowledge at the beginning and end of the activity. This pre/post survey was developed by the DDoE. Results were analyzed to see whether participants reported higher levels of content or pedagogical knowledge after the activity. Third, student writing samples were obtained in December 2014 and again at the end of the school year after participants had piloted lessons from their modules. Participating teachers scored these samples using the state informative writing rubric and shared their ratings with project facilitators, who shared then with the evaluation team. Again results were analyzed to determine whether students showed growth in their writing skills after instruction.

Participant satisfaction and professional development design. Participants' ratings of the design of the PD experience are shown in Table A11, below.

Table A11

Participants' Ratings of Writing Module Development Activity Design

The professional development activity:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
1. Was engaging and kept my attention. ^a	6 (28.6%)	14 (66.7%)	1 (4.8%)	0 (0.0%)	3.24 (0.54)
2. Included a good mix of activities.	2 (9.1%)	16 (72.7%)	4 (18.2%)	0 (0.0%)	2.90 (0.53)
3. Used materials that are relevant to my teaching situation.	11 (50.0%)	10 (45.5%)	1 (4.5%)	0 (0.0%)	3.45 (0.60)
4. Had knowledgeable facilitators/instructors.	9 (40.9%)	13 (59.1%)	0 (0.0%)	0 (0.0%)	3.41 (0.50)
5. Gave me sufficient opportunities to ask questions.	6 (27.3%)	15 (68.2%)	1 (4.6%)	1 (4.6%)	3.22 (0.53)
6. Encouraged meaningful collaboration with colleagues.	16 (72.7%)	5 (22.7%)	1 (4.5%)	0 (0.0%)	3.68 (0.57)
7. Tried to cover too much in the time available. ^a	4 (19.0%)	14 (66.7%)	3 (14.3%)	0 (0.0%)	3.05 (0.59)
8. Modeled effective instruction.	4 (18.2%)	13 (59.1%)	5 (22.7%)	0 (0.0%)	2.95 (0.65)
9. Took place in a space that met our needs. ^a	5 (23.8%)	16 (76.2%)	0 (0.0%)	0 (0.0%)	3.23 (0.44)
10. Addressed a topic that was new to me.	6 (27.3%)	7 (31.8%)	8 (36.4%)	1 (4.5%)	2.81 (0.91)
11. Was worth my time.	10 (45.5%)	9 (40.9%)	3 (13.6%)	0 (0.0%)	3.32 (0.72)

N=22.

^a Missing one response. Percentage is based upon valid responses.

Participants in this activity varied in terms of prior experience (item 10) with just under 60% agreeing the topic was new to them and 40% reporting familiarity. However, prior experience did not appear to be associated with differences in ratings. Participants unanimously agreed that the PD activity had knowledgeable facilitators (item 4). Ratings related to materials and team collaboration (items 3 and 6) were also particularly positive. On the other hand, several participants expressed some concern with the mix of activities and the instruction (items 2 and 8). Group PD sessions were observed to be information-dense, with limited evidence of active learning.

Participants were asked about their favorite aspects of the PD activity, as well as any that need improvement. Themes of responses are shown in Tables A12 and A13, below.

Table A12

Participants' Favorite Parts of the Writing Module Development Activity

Response (Category)	Number	Percentage
Collaborating or working with one's team	9	42.9
Developing a usable/successful module	6	28.6
Learning new skills	5	23.8
Networking or collaborating more broadly	4	19.0

N=21. One respondent skipped this item. Several provided multiple responses.

This activity was highly collaborative, and the largest group of participants wrote about the team aspect: *"I enjoyed working with my team and seeing all of their strengths working together."* Some also mentioned the value of collaborating with a broader network (within a school, or in other districts). Another group of comments focused on the modules produced, calling them *"usable," "meaningful"* and *"innovative."* Finally, some teachers described the development process as a valuable learning experience: *"I learned about my strengths and weaknesses when it came to understanding CCSS, UDL, and performance tasks."*

Table A13

Participants' Suggested Changes to the Writing Module Development Activity

Response (Category)	Number	Percentage
Provide more guidance, clarification, or examples	10	50.0
Provide more meeting time	8	40.0
Start earlier in the school year	4	20.0
Other logistics (scheduling, stipend)	2	10.0

N=20. Two participants skipped this question.

Teachers found developing the module to be a complicated process and many expressed needing more structure and guidance to help them. For instance one recommended *"Giving more specific instructions of how we were to create our Performance Task"* and another stated, *"We needed a model."* Some were unclear about expectations and formatting. Related, several teachers suggest extending the experience (adding more meetings or starting earlier in the year): *"No one wants more meetings but...I would have felt more supported if there were more opportunities to meet and ask questions. I felt like the time together was rushed."* Work time was planned during the meetings but often became truncated. Teachers also described challenges finding time to meet with their teams during the school day, especially if they did not share planning time. Finally, some suggested both teachers and students could benefit from moving this PD forward on the calendar, so there would be more time to teach and build on the module.

Participant learning and instructional change.³ Several data sources help us understand how participants grew during the curriculum development experience. These include participant surveys, participant pre/post questionnaires, and a review of participant work products. Each is discussed below.

Participant surveys. The participant survey included several items related to outcomes.

Table A14

Participants' Ratings of Writing Module Development Learning Outcomes

After participating in this activity:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
12. I will make many adjustments to the module before I teach it again.	7 (31.8%)	10 (45.5%)	5 (22.7%)	0 (0.0%)	3.09 (0.75)
13. I learned new strategies for teaching informative writing.	8 (36.4%)	14 (63.6%)	0 (0.0%)	0 (0.0%)	3.36 (0.49)
14. I plan to share my module with others in my school or district.	11 (50.0%)	7 (31.8%)	4 (18.2%)	0 (0.0%)	3.32 (0.78)
15. The lessons we developed will work well for my students.	10 (45.5%)	10 (45.5%)	2 (9.1%)	0 (0.0%)	3.36 (0.66)
16. I know where to access resources for future curriculum development.	7 (31.8%)	12 (54.5%)	3 (13.6%)	0 (0.0%)	3.18 (0.67)

N=22.

All respondents reported learning new strategies for teaching informative writing (item 13) and almost all felt the modules were appropriate for their students (item 15). They were more mixed on whether they planned to revise their modules and share them with colleagues (items 12 and 14). As discussed further below, the submitted modules were in different stages of refinement.

Most respondents had already taught at least part of their module, and they described mostly positive student reactions (Table A15).

Table A15

Participants' Expectations about Student Reactions to the Writing Modules

Response (Category)	Number	Percentage
Students responded/will respond positively	17	81.0
Students enjoyed/will enjoy the module	7	33.3
Students learned/will learn from the module	7	33.3
Students responded well after modification	1	4.8

N=21. One participant left this item blank. Several included more than one response.

All but one of these comments were positive but they were fairly general. For example: *"Our students enjoyed the unit, they learned a ton, and had fun."* Only one teacher described student resistance (*"the original template was not well received"*) but stated s/he had made *"significant modification."*

Participants were asked what challenges they foresaw to implementing the modules (Table A16).

Table A16

Participants' Expected Challenges to Implementing Writing Modules

Response (Category)	Number	Percentage
Finding time to plan or implement	9	47.4
Positive statement or no challenges	4	21.1
Differentiating module for students	2	10.5
Revising the module	2	10.5
Other (alignment, performance task, sharing with others)	2	10.5

N=19. Three respondents skipped this question.

Teachers foresaw a variety of challenges, with time constraints chief among them. Reflecting on his/her experience this year, one wrote: *"it was very time consuming to plan this unit. It will be very hard to plan something this involved again. Also, the time to teach a unit like this. It took a lot of instruction time. With all we have to teach now it is hard to fit a unit this large into your plans."* Others issues identified include differentiation, alignment, and the need to keep readings/topics up-to-date.

Participant pre- and post-questionnaires. Table A17 shows participants' ratings of their content knowledge both before and after the PD.

Table A17

Teacher's Perception of Content Knowledge Before and After Professional Development in Writing

Content Knowledge Dimension	Deep		Proficient		Adequate		Limited	
	Pre-Survey	Post-Survey	Pre-Survey	Post-Survey	Pre-Survey	Post-Survey	Pre-Survey	Post-Survey
CCSS for grade level	2 (8.7%)	3 (14.3%)	11 (47.8%)	15 (71.4%)	10 (43.5%)	3 (14.3%)	0 (0.0%)	0 (0.0%)
Planning Integrated Units	3 (13.0%)	5 (23.8%)	13 (56.5%)	14 (66.7%)	7 (30.4%)	1 (4.8%)	0 (0.0%)	1 (4.8%)
Six Shifts in ELA/Literacy	1 (4.4%)	3 (14.3%)	1 (4.4%)	12 (57.1%)	18 (78.3%)	5 (23.8%)	3 (13.0%)	1 (4.8%)
Formative Assessments	3 (13.0%)	11 (50.0%)	13 (56.5%)	9 (40.9%)	7 (30.4%)	2 (9.1%)	0 (0.0%)	0 (0.0%)
Universal Design for Learning	0 (0.0%)	2 (9.1%)	5 (21.7%)	13 (59.1%)	7 (30.4%)	5 (22.7%)	11 (47.8%)	2 (9.1%)
CCSS Text-Complexity Matrix	0 (0.0%)	5 (22.7%)	7 (30.4%)	12 (54.6%)	14 (60.9%)	5 (22.7%)	2 (8.7%)	0 (0.0%)
Non-print texts to match CCSS	1 (4.4%)	6 (27.3%)	2 (8.7%)	12 (54.6%)	14 (60.9%)	3 (13.6%)	6 (26.1%)	1 (4.6%)

To examine whether these changes in content knowledge and pedagogy were statistically significant, a Mann-Whitney U test was calculated. In terms of content knowledge, educators reported significantly better understanding of the CCSS after receiving PD ($U=335.5$, $p<.05$). Furthermore, they reported significantly improved knowledge of the Six Shifts in ELA/Literacy ($U=409.50$, $p<.05$), formative assessments ($U=349.5$, $p<.05$), Universal Design for Learning ($U=396.00$, $p<.05$), the CCSS Text-Complexity Matrix ($U=383.00$, $p<.05$), and strategies to find appropriate non-print texts to match CCSS requirements ($U=429.00$, $p<.05$). There was no difference in participants' reports of their understanding of how to plan integrated units ($U=321.5$, $p>.05$).

In addition, participants were asked to rate their pedagogical knowledge on two items. Although the instrument provided different versions of this item for special educators, regular education teachers, and librarians, data were analyzed in the aggregate because of the small numbers (Table 18).

Table A18

Teacher's Perception of Pedagogy Before and After Professional Development in Writing

Pedagogy Dimension	Very Comfortable		Somewhat Comfortable		Somewhat Uncomfortable		Uncomfortable	
	Pre-Survey	Post-Survey ^a	Pre-Survey	Post-Survey ^a	Pre-Survey	Post-Survey ^a	Pre-Survey	Post-Survey ^a
Comfort Level Meeting All Students' Needs	5 (26.3%)	13 (72.2%)	14 (73.7%)	5 (27.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Comfort Level Planning for and Co-Teaching	13 (68.4%)	16 (84.2%)	6 (31.6%)	3 (15.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

N=19.

^a N=18.

In terms of pedagogy, there was a significant change in the improvement of participant comfort level for meeting the needs of all of the students in an inclusion classroom after the PD ($U=290.00$, $p<.05$); however, there were no reported difference with participant comfort level for planning and co-teaching writing instruction ($U=217.00$, $p>.05$).

Participant work products. Evaluators attended the May meeting in which teams presented their modules, and obtained copies of the five final submitted modules. Project staff developed a checklist of criteria for the final modules and shared this with participants (see Appendix 6). The evaluation team adapted this checklist into an instrument and rated to what extent each of 12 criteria was evident in each module. Two evaluators reviewed each module and came to consensus on each rating.

Table A19

Distribution of Criterion Ratings for Informative Writing Modules

Criterion		Meets	Partly Meets	Does Not Meet
1	Planning chart, performance task and lessons aligned	4 (80.0%)	1 (20.0%)	0 (0.0%)
2	Formative assessments aligned/explained in teacher notes	2 (40.0%)	2 (40.0%)	1 (20.0%)
3	Standards and progressions explicitly taught	2 (40.0%)	3 (60.0%)	0 (0.0%)

Criterion		Meets	Partly Meets	Does Not Meet
4	Differentiation embedded throughout lessons	0 (0.0%)	3 (60.0%)	2 (40.0%)
5	Graphic organizers, teaching strategies & vocabulary aligned	5 (100.0%)	0 (0.0%)	0 (0.0%)
6	Format consistent across lessons	3 (60.0%)	2 (40.0%)	0 (0.0%)
7	Pre-requisites listed	4 (80.0%)	0 (0.0%)	1 (20.0%)
8	Appendices and attachments included and labeled	1 (20.0%)	4 (80.0%)	0 (0.0%)
9	Citations included and correct	1 (20.0%)	4 (80.0%)	0 (0.0%)
10	Introduction to module included	2 (40.0%)	0 (0.0%)	3 (60.0%)
11	Presentation included	2 (40.0%)	0 (0.0%)	3 (60.0%)
12	Annotated student work included	4 (80.0%)	0 (0.0%)	1 (20.0%)

N=5.

We made no attempt to “weight” the items on the project checklist, but we note that the first five criteria reflect more complex instructional design issues (alignment, formative assessment, differentiation). Criteria 7 – 12 address formatting and completeness. Given that this activity attempted to create a repertoire of “model units,” these criteria are important but they may be simpler to satisfy.

Ratings of the modules suggest that overall alignment to the CCSS and the learning progressions is strong (criterion 1) and that the organizers, teaching strategies and vocabulary used in each lesson align (criterion 5). These units show efforts to address the CCSS directly. They also generally show effort. Formatting was generally fine in most modules although appendices and citations could have used more attention in some.

However, the modules also show inconsistent evidence of some of the major emphases of this activity. In particular, the expected level of differentiation was not evident in any module. Some modules included meaningful formative assessment and explained how it might be used but others did not. We note that educators reported that they had grown in some of these areas (see above) but the modules they produced do not yet show mastery. These concepts are more difficult for educators to master and more time or guidance may have been necessary.

Student achievement. At least one teacher from each team was expected to pilot at least one of the lessons in the module. As part of this process, they obtained student writing samples both before and after the lessons were taught. Teachers scored student writing samples using the state rubrics for argumentative/opinion writing. Pre- and post-instruction scores were shared with the evaluation team.

The state rubrics included four separate rubrics assessing dimensions of writing. Scores for each dimension ranged from 1 to 4. The total score was obtained by weighting separate rubrics: Reading and Research (x2), Development (x3), Organization (x2), and Language/Conventions (x1). Thus the total score could range from 8 to 32 points.

At the 3rd grade level, the pre and posttest scores of 66 students were included in the analysis (Table A20). A paired-samples *t* test was calculated to compare the mean total pretest score to the mean total posttest score. The mean of the pretest was 15.35 (*sd* = 7.70), and the mean on the posttest was 22.30 (*sd* = 6.32). A significant increase from pretest to posttest was found ($t(df) = -7.14, p < .001$). There were also significant differences between pretests and posttests for all four subdomains (Reading/ Research, Development, Organization, and Language/Conventions). The domain demonstrating the most improvement was Development.

Table A20
Pre and Post Instruction Rubric Scores for 3rd Grade Students

Rubric	Mean Pre Score	Mean Post Score	Mean Difference	<i>t</i>
Reading/Research	1.98	2.86	.88	5.97*
Development	1.89	2.92	1.03	6.96*
Organization	1.86	2.59	.73	5.65*
Language/Conventions	1.97	2.62	.65	5.56*
Total Score	15.35	22.30	6.96	7.14*

*Statistically significant ($p < .001$).

In grade 5, the pre and posttest scores for 63 students were included in the analysis (Table A21). A paired-samples *t* test was calculated to compare the mean total pretest score to the mean total posttest score. The mean of the pretest was 15.37 (*sd* = 5.77), and the mean on the posttest was 19.71 (*sd* = 7.96). A significant increase from pretest to posttest was found ($t(62) = 6.33, p < .001$). There were also significant differences between pretests and posttests for all four subdomains. Again the domain demonstrating the most improvement was Development.

Table A21

Pre and Post Instruction Rubric Scores for 5th Grade Students

Rubric	Mean Pre Score	Mean Post Score	Mean Difference	t
Reading/Research	1.97	2.56	.587	5.86*
Development	1.89	2.49	.603	6.03*
Organization	1.90	2.35	.444	5.11*
Language/Conventions	1.95	2.35	.476	4.98*
Total Score	15.37	19.71	4.40	6.33*

*Statistically significant ($p < .001$).

In grade 9, the pre and posttest scores for 61 students were included in the analysis (Table A22). A paired-samples t test was calculated to compare the mean total pretest score to the mean total posttest score. The mean of the pretest was 15.75 ($sd = 4.77$), and the mean on the posttest was 21.68 ($sd = 4.97$). A significant increase from pretest to posttest was found ($t(60) = 9.79$, $p < .001$). There were also significant differences between pretests and posttests for all four subdomains. In this grade level, the greatest improvement was seen in Reading/Research.

Table A22

Pre and Post Instruction Rubric Scores for 9th Grade Students

Rubric	Mean Pre Score	Mean Post Score	Mean Difference	t
Reading/Research	3.73	5.90	2.17	10.22*
Development	6.10	7.82	1.72	5.79*
Organization	3.54	5.02	1.48	7.80*
Language/Conventions	2.39	2.94	.56	4.65*
Total Score	15.75	21.68	5.93	9.79*

*Statistically significant ($p < .001$).

Looking across all the data for this activity, we see that on average students in all three grade levels scored higher on their writing pieces after instruction. After participating in the activity, educators reported higher levels of content knowledge and pedagogical comfort in some, but not all, of the target areas. A review of the modules themselves suggests that considerable work went into them but some areas, especially differentiation, remain uneven. Finally, participant responses to the experience were overall positive but suggest the need for more guidance in the future.

Summer Writing Institute

Introduction. *Teaching Non-Fiction Writing with Style* was a Delaware Writing Project (DWP) summer institute for ELA, science and social studies teachers in K-12. It took place on six afternoons in July at UD and was facilitated by two teachers who are Delaware Writing Project (DWP) alumni. The 17 participants included elementary, middle and high school teachers as well as instructional coaches. This

institute focused on the non-fiction writing emphasized by the CCSS, especially informative/explanatory texts (Standard 2). Through a variety of activities and mentor texts, this institute aimed to help teachers see that such writing need not be “dry and boring.” According to survey results (see below) slightly more than half of the participants found this topic new to them. At least four had attended a Title II writing institute in a previous summer.

This hybrid institute combined face-to-face PD with independent work and posts on Schoology, a Learning Management System now widely available in the state. Participants were asked to bring their laptops. The institute incorporated a textbook (Culham, 2014) as well as selected articles. Consistent with the principles of the DWP, the institute emphasized that educators learn about writing by writing. Every day included warm-up writing prompts and writing activities, many of which were completed on Schoology. Facilitators also modeled some non-fiction writing activities and engaged participants in discussing how these could be adapted for their classrooms. On one day, two UD English department faculty members attended the institute to share ideas for context-specific writing assignments.

Although originally the institute was to be geared towards informative/explanatory writing, in practice it looked more broadly at non-fiction writing. Throughout the institute, participants developed a non-fiction writing lesson plan. They shared a draft of these plans with their peers on the last day, then submitted a final version shortly afterwards. Sixteen of the 17 participants completed this requirement.

In the following section, we summarize evaluation results that include participant responses to the survey administered at the end of the institute, evaluators’ observations, a review of the Schoology site, and an analysis of the participants’ submitted lesson plans.

Participant satisfaction and professional development design. Participants were asked to rate many aspects of the institute’s design and their overall satisfaction (Table A23).

Table A23
Participants’ Ratings of Summer Writing Institute Design

The institute:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
1. Was engaging and kept my attention.	6 (37.5%)	8 (50.0%)	2 (12.5%)	0 (0.0%)	3.25 (0.68)
2. Included a good mix of activities.	6 (37.5%)	8 (50.0%)	2 (12.5%)	0 (0.0%)	3.25 (0.68)
3. Used materials that are relevant to my teaching situation.	9 (56.3%)	7 (43.8%)	0 (0.0%)	0 (0.0%)	3.56 (0.51)
4. Had knowledgeable facilitators/instructors.	12 (75.0%)	4 (25.0%)	0 (0.0%)	0 (0.0%)	3.75 (0.45)
5. Gave me sufficient opportunities to ask questions.	13 (81.3%)	3 (18.8%)	0 (0.0%)	0 (0.0%)	3.81 (0.40)

The institute:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
6. Encouraged meaningful collaboration with colleagues.	12 (75.0%)	3 (18.8%)	1 (6.3%)	0 (0.0%)	3.69 (0.60)
7. Tried to cover too much in the time available.	0 (0.0%)	0 (0.0%)	10 (62.5%)	6 (37.5%)	1.63 (0.50)
8. Modeled effective instruction.	4 (25.0%)	11 (68.8%)	1 (6.3%)	0 (0.0%)	3.19 (0.54)
9. Took place in a space that met our needs.	10 (62.5%)	6 (37.5%)	0 (0.0%)	0 (0.0%)	3.63 (0.50)
10. Addressed a topic that was new to me.	2 (12.5%)	7 (43.8%)	7 (43.8%)	0 (0.0%)	2.69 (.70)
11. Was worth my time.	7 (43.8%)	8 (50.0%)	1 (6.3%)	0 (0.0%)	3.38 (0.62)

N=16.

These responses suggest generally high satisfaction, although two participants were more critical about their experience and were responsible for most of the “disagree” ratings. In general, participants appreciated many aspects of the institute including the materials, the facilitation, and the opportunity to ask questions and collaborate. From the introductions on the first day, it was apparent that participants had different reasons for attending the institute and various levels of previous experience (also shown in item 10). Observation notes show that most participants were engaged but a few were off-task at any one time. Sometimes this could be accounted for by role. For example, high school teachers appeared more engaged in the presentation by the UD faculty members than teachers of elementary or middle.

Unusual for a Title II activity, participants unanimously found the agenda realistic for the time available (item 7). On the other hand, ratings and open-ended comments suggest that some participants wanted more engaging activities and a greater mix of them (items 1 and 2). On some days, the activities were slow to get started or a facilitator was late due to extenuating circumstances. Ratings for whether the workshop “modeled effective instruction” were somewhat mixed.

Participants were asked to identify their favorite aspects of the institute and to suggest changes. See Tables A24 and A25, below.

Table A24

Participants’ Favorite Parts of the Summer Writing Institute

Response (Category)	Number	Percentage
Collaboration with fellow educators	10	62.5
Writing strategies modeled	3	18.8
Enthusiasm for further study/follow up	2	12.5
Other (workshop environment, flexibility of final assignment)	2	12.5

N=16. One included more than one response.

The majority of responses referred to collaboration: *“the opportunity to engage in a meaningful, positive peer communication.”* A smaller group described the benefits of modeling: *“seeing how the presenters model the strategies is so helpful in guiding my own decisions.”* A few other comments were more difficult to categorize.

Table A25

Participant Suggestions for Changes to the Summer Writing Institute

Response (Category)	Number	Percentage
Improve activities or include more of them	5	33.3
Feature more modeling/instruction of how to teach writing	3	20.0
Nothing or only positive comments	3	20.0
Make it longer or have it more often	3	20.0
Use more or different resources/texts	2	13.3
Other specific suggestion	1	6.7

N=15. One participant skipped this question.

Over half the comments suggested improvements to the activities or instruction. Some teachers requested more direct instruction: *“I would have liked to have been taught how to teach non-fiction writing. I feel that I learned a lot of strategies but still struggle to actually teach this genre of writing.”* Several teachers echoed this idea, stating that the activities needed to be more “specific” or “organized.” Others suggested making the activities more hands-on or collaborative. One pointed out that the institute did not fully take advantage of Schoology. While teachers posted their writing reflections and responses to the articles online, there was little evidence of dialogue among participants or between participants and facilitators.

Participant learning. Data come from participant surveys and a review of their work products.

Participant surveys. The participant survey included several items to measure outcomes.

Table A26

Participants’ Ratings of Summer Writing Institute Learning Outcomes

After participating in the institute:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
12. I learned new approaches for teaching students informative writing.	7 (43.8%)	7 (43.8%)	2 (12.5%)	0 (0.0%)	3.31 (0.70)
13. I better understand how to align my lessons to the CCSS.	4 (25.0%)	8 (50.0%)	4 (25.0%)	0 (0.0%)	3.00 (0.73)
14. I have seen examples of activities that would work well for students.	8 (50.0%)	8 (50.0%)	0 (0.0%)	0 (0.0%)	3.50 (0.52)
15. I received helpful feedback on my lesson plan as I developed it.	5 (31.3%)	9 (56.3%)	2 (12.5%)	0 (0.0%)	3.19 (0.66)

After participating in the institute:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
16. Lack of resources at my school will make the strategies we covered difficult to implement.	0 (0.0%)	2 (12.5%)	11 (68.8%)	3 (18.8%)	1.94 (0.57)
17. I feel confident in my ability to implement ideas and practices from the workshop in the 2015-16 school year.	9 (56.3%)	7 (43.8%)	0 (0.0%)	0 (0.0%)	3.56 (0.51)
18. My writing lesson plan needs more work before it is ready to be used.	0 (0.0%)	9 (56.3%)	7 (43.8%)	0 (0.0%)	2.56 (0.51)
19. I would like to participate in follow-up PD to support my writing instruction.	8 (50.0%)	8 (50.0%)	0 (0.0%)	0 (0.0%)	3.50 (0.52)

N=16.

Some of these ratings were unanimously positive. All participants reported they had seen relevant example activities (item 14), felt confident in their ability to implement workshop content in instruction (item 17), and wanted to participate in follow-up PD (item 19).

Other items yield mixed responses. Participants were divided on whether their lesson plans need development (item 18) – inconsistency that was also evident in the document review (see below). Two respondents disagreed that the institute taught them new strategies for writing instruction; both reported having prior experience with the topic. While the schedule provided many opportunities for feedback, response to the usefulness of that feedback were lukewarm (item 15). During work time, facilitators were observed in conversation with participants but occasionally those conversations appeared off task. On the final lesson plans submitted on Schoology, some but not all included wr

To better understand participants' instructional plans, they were asked how they planned to embed non-fiction writing in their curriculum. Themes of responses are in Table A27.

Table A27

Participants' Plans to Implement Writing Strategies into their Curricula

Response (Category)	Number	Percentage
Use mentor texts	8	50.0
Use activity or resource from summer institute in instruction	7	43.8
Ask students to write more often/in a variety of formats or contexts	6	37.5
Incorporate technology or multimedia	4	25.0
Write across curricular areas	3	18.8
Build on existing practices	2	12.5

N=16. Many participants included multiple responses.

Reflecting one of the main emphases of the institute, half the respondents stated they would incorporate mentor texts in their instruction. Seven planned to use activities or resources from the institute; some of these comments were general and others specific. For instance one identified an

activity where the group brainstormed “18 writing pieces from one topic” as something she “could share w/staff.” The third group of responses described changing the way they had students write, using more frequent and multiple opportunities: “I also plan to do smaller activities and quick writes, instead of focusing on big projects.” Writing regularly was another main message of the institute.

Participants were also asked to report how they felt their students would react to the writing strategies taught in the institute (Table A28) and what challenges they expected (Table A29).

Table A28

Participants’ Expectations about Student Reactions to the Institute’s Writing Strategies

Response (Category)	Number	Percentage
Students would respond positively (general statement)	6	40.0
Students would be engaged or excited	6	40.0
A specific activity or strategy would work well for students	3	20.0
Strategies provide choices or differentiation for students	3	20.0
Strategies will require some adjustment	1	6.7

N=15. One participant skipped this question. Several included more than one response.

Almost all of these comments were positive. Participants anticipated that students would enjoy the activities and find them engaging, novel, and manageable. For example, one participant wrote: “I think that the instructors spoke a lot about student choice and student excitement. Kids would love many of the strategies.” Another enthused, “My students are going to see themselves as writers. I know this!”

Table A29

Participants’ Expected Challenges to Implementation of Writing Strategies

Response (Category)	Number	Percentage
Time constraints	4	30.8
None or N/A	4	30.8
Buy-in from others	3	23.1
Student needs	2	15.4

N=16. One included more than one response.

Several themes emerged as possible challenges to implementation. The most common, as one participant wrote, was “Time! Always!” Other teachers were concerned about getting colleagues or teammates to agree with the plans. Finally, two respondents identified challenges around “various instructional levels” or helping students use prior knowledge in the writing strategies.

Participant work products. Sixteen participants (94.1%) submitted lesson plans at the end of the institute via Schoology. These were reviewed with a rubric developed by institute facilitators and shared with participants. It is available in Appendix 7. One dimension, level of lesson development, was added

by evaluators. Two evaluators rated each lesson plan and came to consensus on the ratings. The distribution of ratings for the 16 lesson plans on each rubric is shown in Table A30, below.

Table A30
Distribution of Rubric Ratings for Participants' Draft Non-Fiction Writing Lesson Plans

Rubric	Outstanding	Good	Acceptable	Needs Work
Standards	0 (0.0%)	1 (6.3%)	10 (62.5%)	5 (31.3%)
21 st Century Literacy	2 (12.5%)	4 (25.0%)	10 (62.5%)	0 (0.0%)
Differentiation/Scaffolding	2 (12.5%)	2 (12.5%)	8 (50.0%)	4 (25.0%)
Exemplary Teaching of Writing Practices	2 (12.5%)	6 (37.5%)	8 (50.0%)	0 (0.0%)
Mentor Text(s)	3 (18.8%)	10 (62.5%)	3 (18.8%)	0 (0.0%)
Extended Thinking	3 (18.8%)	2 (12.5%)	3 (18.8%)	8 (50.0%)
Final Writing Task	2 (12.5%)	8 (50.0%)	4 (25.0%)	2 (12.5%)
Level of Development	1 (6.3%)	8 (50.0%)	6 (37.5%)	1 (6.3%)

N=16.

This table reflects some of the foci of the institute. For example, the institute placed a great deal of emphasis on mentor texts and on research-based writing strategies more generally. All plans demonstrated at least acceptable evidence of these two criteria and at least half were rated “good” or “outstanding.” On the other hand, these data also reflect overall inconsistent quality of lesson plans, and this was also evident in the final presentations. The level of differentiation/scaffolding and extended thinking in the plans was notably uneven. In terms of overall development, nine plans were rated “good” or “outstanding” and seven were “acceptable” or “needs work.”

The intention was for facilitators to provide written feedback on the submitted lesson plans. Of the sixteen plans submitted on and available through Schoology, ten (62.5%) include feedback and six (37.5%) do not. No information is available about why they do not.

Instructional change and student achievement. Due to the timing of this activity, no data were available for these evaluation questions. Follow-up activities have been proposed that would allow ongoing support for participants as they teach, assess, and revise their plans. All participants from the summer expressed interest in remaining involved.

Social Studies Webinar Series

Introduction. This activity was open to teachers in grades K-12 as well as school and district administrators responsible for overseeing social studies instruction. It aimed to build participants' content and pedagogical content knowledge about social studies, and about the alignment of social studies to the ELA CCSS. Activities took place exclusively online. The series included ten hour-long webinars between December 2014 and May 2015, focused on key topics in social studies instruction, the CCSS, and the Smarter Balanced assessment. These webinars were organized and facilitated by a staff member from UD's Professional Development Center for Educators; some sessions also included guest presenters. Recruitment occurred through statewide emails. Participants registered for each webinar separately. Once registered, they could attend the webinar live and/or access a recording later. They were also encouraged to watch it with or disseminate materials to colleagues. The intent of this professional development design was to provide flexibility and ease of access. For instance, it was hoped that teachers might watch/discuss a webinar during a PLC.

Participation data for the webinars was obtained from project staff. In total, 87 educators in elementary, middle and high schools registered for at least one webinar.⁴ Of these, 29 (33.3%) attended at least one webinar live. It is not possible to determine how many of the remaining 58 later accessed the recordings although participant surveys (see below) indicate this did occur at least some of the time. On the other hand, surveys also indicate that some registrants never viewed any webinar. Participation levels varied. Of the 87 registrants, 65 (74.7%) registered only for a single webinar in the ten-webinar series and twelve (13.8%) registered for just two. Small numbers registered for three or four (5.7%) and five or more (4.6%) webinars. There were two "regulars" who registered for more than half of the webinars. No one registered for all ten.

Webinars in the series attracted different levels of registration and live attendance (Table A31).

Table A31

Social Studies Webinars Dates, Topics, and Number of Attending and Registered

Date	Topic	Number Attending	Number Registered
December 4, 2014	Overview of Delaware social studies standards and instructional resources	10	12
December 18, 2014	Introduction to social studies in Delaware	4	1
January 7, 2015	Overview of Smarter Balanced performance task design and what it might look like for social studies	10	20
January 22, 2015	Overview of Smarter Balanced ELA/literacy assessment: Computer Adaptive Test	7	6
February 12, 2015	Disciplinary literacy Part I	6	4
February 19, 2015	Disciplinary literacy Part II	5	2

⁴ This number (and all subsequent registration/attendance data) does not include staff members from UD or DDoE.

Date	Topic	Number Attending	Number Registered
March 31, 2015	Problematic prior learning in history	4	13
April 30, 2015	Evidence-based social studies strategies that engage, I	0	11
May 20, 2015	Evidence-based social studies strategies that engage, II	5	4
May 27, 2015	Academic vocabulary in social studies	5	4

With these data as context, we turn now to findings from the participant survey. Due to the online format for this activity, our usual data collection procedures were adapted. A survey invitation and link to a Qualtrics survey were sent to each individual who had registered for at least one webinar (N=87). From this group, we obtained 31 responses (total response rate, 35.6%). Five responses (16.1%) indicate that the individual never viewed any webinar.⁵ They cited barriers such as “time,” “other things going on” and “computer issues.” The rest of this section includes responses from the 25 survey respondents who attended at least one webinar. It also integrates evaluator observations.

Respondents included elementary school teachers (22%); middle school teachers (44%); high school teachers (16%); administrators from the school, district or state (15%); and an instructional coach (3%). As expected from the project-provided attendance data, these respondents reported various methods of accessing the webinars. Twenty (80%) attended at least one webinar live; fifteen (60%) watched at least one recording; and nine (36%) registered with the intent of watching a webinar later. Respondents were asked whether and how they disseminated webinar materials to others (Table A32).

Table A32
Dissemination of Webinar Materials

Method of dissemination	Number	Percentage
I sent colleagues the webinar recording or registration link	9	34.6
I talked with colleagues about the information in the webinars	13	50.0
I viewed the webinar or the recording together with colleagues	7	26.9
I did not share webinars or information from webinars with colleagues	5	19.2

n=26. Percents do not add to 100, participants were able to select multiple responses.

The majority of respondents used at least one of these methods of dissemination. Over a third had passed on the webinar materials directly and half discussed it indirectly with colleagues. There is some evidence that groups of educators watched the webinar together (live or recording) though this does not appear prevalent.

⁵ In addition, evaluators received three email responses to the survey invitation. In each case, the respondent stated s/he never attended any webinars. Reasons cited included connection issues and competing demands of an administrative position.

Survey respondents varied in their prior exposure to social studies PD; ten (40%) had received more than a week of PD in social studies before these webinars and the same proportion had received a little (two days or less). The remaining participants had no (16%) or some (4%) prior PD related to social studies. To this end, facilitators presented information about how Smarter Balanced assessments are formatted, in addition to presenting example assessments and discussing social studies and history pedagogy. Participants spent the majority of the webinar sessions listening to informative PowerPoint presentations by the facilitators and periodically responding to poll or text entry questions. Presenters paused to allow for participant-initiated questions (either over the mic or in the text box) though rich dialogue did not typically occur.

Participant satisfaction and professional development design. Participant ratings for the items related to professional development quality and overall satisfaction are shown in Table A33.

Table A33

Participant Ratings of the Social Studies Webinar Series Design

The professional development:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
1. Was engaging and kept my attention. ^a	6 (24.0%)	18 (72.0%)	1 (4.0%)	0 (0.0%)	3.2 (0.50)
2. Included a good mix of activities.	5 (20.0%)	18 (72.0%)	2 (8.0%)	0 (0.0%)	3.12 (0.53)
3. Used materials that are relevant to my teaching situation.	8 (32.0%)	15 (60.0%)	2 (8.0%)	0 (0.0%)	3.24 (0.60)
4. Had knowledgeable facilitators/instructors.	15 (60.0%)	10 (40.0%)	0 (0.0%)	0 (0.0%)	3.60 (0.50)
5. Gave me sufficient opportunities to ask questions.	5 (20.0%)	18 (72.0%)	1 (4.0%)	1 (4.0%)	3.08 (0.64)
6. Encouraged meaningful collaboration with colleagues.	3 (12.0%)	19 (76.0%)	3 (12.0%)	0 (0.0%)	3.00 (0.50)
7. Tried to cover too much in the time available. ^a	0 (0.0%)	9 (37.5%)	15 (62.5%)	7 (29.2%)	2.67 (0.49)
8. Modeled effective instruction.	3 (12.0%)	20 (80.0%)	1 (4.0%)	1 (4.0%)	3.00 (0.58)
9. Addressed a topic that was new to me. ^a	5 (20.8%)	15 (62.5%)	2 (8.3%)	2 (8.3%)	2.96 (0.81)
10. Was worth my time. ^a	7 (29.2%)	16 (66.7%)	1 (4.2%)	0 (0.0%)	3.25 (0.53)

N=25.

^a N=24.

Survey respondents were generally satisfied. Almost all (95.87%) agreed/strongly agreed that the webinar series was worth their time. They unanimously found the facilitators knowledgeable. Webinar observations confirm that a large amount of information about social studies standards, content, and instructional strategies was shared, and participants appeared to find it useful.

On the other hand, these ratings suggest some areas for improvement. While candidates mostly responded positively to statements about collaboration, opportunities for questions, and effective instruction, they did not often choose the highest rating. Observations suggest the webinars focused on delivering content; attempts at dialogue were made, but were mostly not successful.

Participants were asked about their favorite aspects of the webinar series as well as any they would recommend changing (Tables A34 and A35).

Table A34

Participants' Favorite Parts of the Social Studies Webinar Series

Response (Category)	Number	Percentage
Convenience of having taped recordings	6	35.3
Relevant webinar content	5	29.4
Format of online sessions	3	17.6
Knowledgeable presenters	2	11.8

N=17. One included more than one response. Eight survey respondents left this question blank.

Table A35

Participant Suggestions for Changes to the Social Studies Webinar Series

Response (Category)	Number	Percentage
No suggestions or general positive comment	5	41.7
Adjust logistics and timing of webinar sessions	3	25.0
Alter presentation of content	3	25.0
Fix technology issues	2	16.7

N=12. One participant provided multiple responses. Thirteen survey respondents left this item blank.

Participant comments indicate that participants particularly enjoyed the convenience of the online format and the relevant webinar content. Regarding the content, one participant wrote: *"I liked that the focus was on Social Studies! I feel like education is losing focus on the subject."* Attendees also suggested things they would change about the webinar series. The majority of suggestions included shifts to logistics and content presentation (50.0%). For example, one participant suggested *"breaking it into testing groups and gear these mainly to the specific grades that test"* while another suggested *"allowing teachers to view PD slides before the event to see if it will be relevant."* Several respondents suggested that starting webinars later in the afternoon would work better for elementary teachers.

Due to the format of this PD activity, unique among Title II activities, we added several items to this survey related to the online format and webinar technology. Responses to survey items are found in Table A36 and participants comments related to technology are presented in Table A37.

Table A36

Participant Ratings of Online Format of Social Studies Webinars

Indicate your agreement with each statement	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
1. The webinar technology was user-friendly.	5 (20.0%)	17 (68.0%)	2 (8.0%)	1 (4.0%)	3.04 (0.68)
2. I had access to technology (e.g., mic, headphones) to allow me to participate fully.	2 (8.0%)	18 (72.0%)	3 (12.0%)	2 (8.0%)	2.88 (0.71)
3. Webinars took place at a convenient time for me.	1 (4.0%)	17 (68.0%)	6 (24.0%)	1 (4.0%)	2.72 (0.61)
4. Webinars were a good format for this PD. ^a	8 (33.3%)	16 (66.7%)	0 (0.0%)	0 (0.0%)	3.00 (0.48)
5. Overall the online format encouraged me to participate. ^b	7 (30.4%)	13 (56.5%)	3 (13.0%)	0 (0.0%)	3.17 (0.65)

N=25.

^aN=24.

^bN=23.

Table A37

Participant Responses to Advantages and Disadvantages of Online Format for Social Studies Webinars

Response	Number	Percentage
Advantages		
Convenience of watching recorded sessions on user's own time	5	38.5
General or miscellaneous positive comment about the PD format	4	30.8
Online format allowed participants to watch from a convenient location	2	15.4
Disadvantages		
Technology issues inhibited webinar participation	4	30.8
General or miscellaneous negative comments about the PD format	2	15.4

n=13. Participants provided multiple responses.

In general, participants were satisfied with the format of the Social Studies webinars. All agreed or strongly agreed that webinars were a good format for this professional development (item 4). Participants also generally found the webinar technology user-friendly (item 1), had access to necessary technology (item 2), and stated that the online format encouraged them to participate (item 5). Open-ended comments also reflect positive views about the convenience of the webinar format. One

participant also reported being “able to include several colleagues who would not normally have participated in social studies training.”

Despite the enthusiasm about its convenience, participants also indicated that there were technology issues associated with the online format. One participant stated “we have to get muting under control. There was too much background noise” while others stated they “had no audio” during some of the webinars. Evaluators observed the same difficulties. In addition, the recording feature did not work on at least two occasions, compromising registrants’ ability to access or share the webinar. Finally, it is important to note that a few individuals cited technological issues for why they registered for but were not able to attend.

Participant learning. The participant survey included several items to measure personal and institute outcomes (Tables A38 and A39).

Table A38

Participants’ Ratings of Learning Gains in the Social Studies Webinar Series

How much did participating in this professional development increase your understanding of the following topics?	Large Increase	Some Increase	Little Increase	No Increase	Did not participate in a webinar in this topic
1. Common Core State Standards. ^a	5 (20.8%)	14 (56.0%)	2 (8.3%)	2 (8.3%)	1 (4.2%)
2. Smarter Balanced Assessment. ^a	6 (25.0%)	9 (37.5%)	4 (16.7%)	1 (4.2%)	4 (16.7%)
3. Disciplinary literacy in social studies. ^a	4 (16.7%)	15 (62.5%)	3 (12.5%)	0 (0.0%)	2 (8.3%)
4. Academic vocabulary in social studies	3 (12.0%)	15 (60.0%)	2 (8.0%)	0 (0.0%)	5 (20.0%)
5. Problematic prior learning in social studies	2 (8.0%)	6 (24.0%)	7 (28.0%)	1 (4.0%)	9 (36.0%)
6. Strategies for engaging students in social studies	4 (16.0%)	12 (48.0%)	2 (8.0%)	0 (0.0%)	7 (28.0%)

N=25.

^aN=24.

Sessions on the CCSS and Smarter Balanced were more popularly attended and more positively received relative to other webinar sessions. In addition many participants (79.2%) thought the disciplinary literacy in social studies session had a large or moderate increase to their understanding of the topic.

Table A39

Participant Ratings of Social Studies Webinar Learning Outcomes

Please indicate your agreement with the following:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
1. I have been able to use what I learned in this PD.	2 (8.0%)	22 (88.0%)	1 (4.0%)	0 (0.0%)	3.04 (0.35)
2. Participating in this PD had a positive impact on my instructional or leadership practices.	2 (8.0%)	20 (80.0%)	3 (12.0%)	0 (0.0%)	2.96 (0.45)
3. Participating in this PD had a positive impact on my student learning. ^a	2 (8.7%)	18 (78.3%)	3 (13.0%)	0 (0.0%)	2.96 (0.47)

n=25.

^a n=23

While most of these responses are positive, teachers tended to “agree” rather than “strongly agree” with these statements about application. Only five participants provided open-ended comments to give more information about impacts on teaching and learning. Due to the small response size, comments were not coded thematically. Similar to survey ratings, two of these participants commented on the usefulness of the discussion about how CCSS connects to social studies. The other three comments did not give specific examples of how the sessions impacted instruction and student learning.

Participants were also asked what challenges they experienced or anticipate in implementing what they have learned in the professional development sessions. Twelve participants responded to this item, however the majority of responses did not address the question or indicated there were no challenges to implementing instruction (66.7%). Two participants anticipated barriers due to limited instructional time, one cited technology, and the fourth cited “performance tasks” but did not explain.

Instructional change and student achievement. Participants were invited to submit instructional artifacts and/or student work to show the impact of the webinars and were offered professional development hours credit if they did so. Only one participant submitted an artifact. Due to the low numbers, this was not analyzed.

Responsive Classroom

Introduction. *Responsive Classroom* is a research-based approach to teaching students the social curriculum and building classroom community. Its goal is to support safe, engaging environments where learning can happen. The Massachusetts-based Center for Responsive Schools (CRS) is the sole source provider for *Responsive Classroom*. Through the Title II grant, UD entered into a partnership to provide *Responsive Classroom* PD, supported by individual coaching, to multiple elementary schools in Delaware that had identified the need to improve classroom and school climate.

This was the largest Title II activity as measured by total enrollment, and the only one to use a school wide training approach. UD initially planned to work with eight schools: three in the Colonial School District (New Castle, Pleasantville, and Southern), two in the Red Clay Consolidated School District (Shortlidge and Warner), one in the Christina School District (Oberle) and two charter schools (East Side and Las Américas ASPIRA). All of these schools planned to offer a combination of *Responsive Classroom* PD sessions and follow-up coaching. With consultation from project staff, administrators in each school chose a *Responsive Classroom* “kit” to focus on (Teacher Language, Morning Meeting, or Discipline). The structure and schedule of activities was at the discretion of the school; some arranged full-day trainings, others used more frequent, shorter sessions. In addition, schools varied in the balance of group PD and coaching provided. The total time allocated to each school through the grant for PD and coaching was 42 hours.

Ultimately some of these partnerships proved more robust than others. PD and coaching occurred as planned in the three Colonial Schools and both charter schools. Some but not all of the intended activities occurred at Shortlidge and Warner.⁶ After a few initial sessions, Oberle did not respond to multiple efforts to arrange activities and eventually support from this school was re-allocated to others. Finally, an unexpected opportunity to extend training in the Colonial schools emerged, as that district required teachers to participate in PD to make up for snow days. An after-school *Responsive Classroom* series was developed and supported through Title II.

In addition to the school partnerships, the Title II grant supported three *Responsive Classroom* institutes in summer 2015. The first, *Teaching Discipline in the Classroom*, took place on five consecutive mornings in July, and was free of cost to participants. Two sessions of the week long *Responsive Classroom Courses* were also offered. This introduces participants to all aspects of the *Responsive Classroom* approach to the social curriculum. Due to support from Title II, the cost of attendance was \$500 instead of the regular price of \$725.

A team of four UD faculty and staff members collaborated on these *Responsive Classroom* activities. One served as a project director, led coordination with the schools, and conducted some PD sessions. A second, certified as a *Responsive Classroom* facilitator, led most of the PD and all of the coaching. Two other staff members provided support and assisted with the training.

Evaluation activities for *Responsive Classroom* included observing at least one session at each school that conducted PD as well as each summer institute. We administered a satisfaction survey to participants at the end of each activity (i.e., the last day of the summer institutes, or the last scheduled PD session in each partnership school).⁷ To learn more about the coaching experience, we conducted interviews with teachers who participated in at least one coaching session. The sample included six

⁶ Both of these schools were designated as Priority Schools after they had already agreed to Title II/Responsive Classroom partnerships. Priority Schools planning took priority. Still, individual coaching was offered in both of these schools.

⁷ Because Oberle, Shortlidge and Warner did not complete the PD sessions, they were not included in the participant surveys.

teachers from four partnership schools.⁸ We interviewed administrators in five partnership schools and collected school climate data from these schools.

The following sections present the complete *Responsive Classroom* data. First we present data about participant characteristics, then survey data aggregated across five partnership schools⁹ and three summer institutes. We also present some disaggregated analyses. Perspectives from coaching participant and administrator interviews are integrated where relevant. Finally we present comparative school climate data for 2013-14 and 2014-15.

Participant characteristics. Table A40 shows the number of *Responsive Classroom* surveys received from each partnership school and each summer institute, i.e. the composition of the data set.

Table A40

Number and Percentage of Responsive Classroom Surveys Received by School or Institute

School or Institute	Number	Percentage
Las Américas ASPIRA Academy	40	13.7
East Side Charter School	46	15.8
New Castle Elementary School	35	12.0
Pleasantville Elementary School	33	11.3
Southern Elementary School	53	18.2
Discipline in the Classroom	31	10.6
Responsive Classroom Course (June)	25	8.6
Responsive Classroom Course (July)	29	9.9

N=292.

The majority of participants (61.8%) were elementary school teachers, though 27 (9.3%) reported teaching grades 5-6, 9 (3.1%) taught grades 7-8 and 71 (24.5%) worked with multiple grade levels. In terms of role, just over half (54%) were grade-level classroom teachers and a quarter (25.7%) were special education teachers. The remainder (20.3%) worked with whole school populations, as specialists, administrators, or Related Arts teachers.

Participants were asked how much prior training they had had with *Responsive Classroom* prior to the Title II activity. The majority had participated in either a brief (one day or less) session (29.3%) or a multi-day training (29.0%) but 41.7% reported that they were new to the approach.

Participant satisfaction and professional development design. Table A41 displays all participant ratings related to satisfaction and design for the professional development sessions. Feedback on *coaching* is reported separately in a subsequent section.

⁸ Project records included 48 names of teachers who received coaching. We sampled from this list, focusing on teachers who had participated in two or more coaching sessions, except in schools where all teachers only had a single coaching session. Ultimately we obtained six interviews, representing four of the partnership schools. This sample has limitations (small number, not all schools represented).

⁹ Summaries of each individual school's data were shared with project staff as they became available.

Table A41

Participant Ratings of the Responsive Classroom Workshop Design, Aggregated

The professional development:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
1. Was engaging and kept my attention. (N=290)	124 (42.8%)	96 (33.1%)	51 (17.6%)	19 (6.6%)	3.12 (.92)
2. Included a good mix of activities. (N=292)	135 (46.2%)	115 (39.4%)	37 (12.7%)	5 (1.7%)	3.30 (.75)
3. Used materials that are relevant to my teaching situation. (N=291)	120 (41.2%)	126 (43.3%)	37 (12.7%)	8 (2.7%)	3.23 (.77)
4. Had knowledgeable facilitators/instructors. (N=292)	169 (57.9%)	109 (37.3%)	10 (3.4%)	4 (1.4%)	3.52 (.63)
5. Gave me sufficient opportunities to ask questions. (N=290)	140 (48.3%)	121 (41.7%)	25 (8.6%)	4 (1.4%)	3.37 (.70)
6. Encouraged meaningful collaboration with colleagues. (N=291)	153 (52.6%)	108 (37.1%)	25 (8.6%)	5 (1.7%)	3.41 (.72)
7. Tried to cover too much in the time available. (N=292)	15 (5.1%)	68 (23.3%)	173 (59.2%)	36 (12.3%)	2.21 (.72)
8. Modeled effective instruction. (N=290)	135 (46.6%)	117 (40.3%)	33 (11.4%)	5 (1.7%)	3.32 (.74)
9. Addressed a topic that was new to me. (N=289)	68 (23.5%)	124 (42.9%)	73 (25.3%)	24 (8.3%)	2.82 (.89)
10. Was worth my time. (N=288)	129 (44.8%)	94 (32.6%)	41 (14.2%)	24 (8.3%)	3.14 (.95)

Number of respondents varies by item and is noted next to each item.

The majority of these ratings were positive although on certain items we see quite wide distributions. Ninety percent or more of respondents agreed/strongly agreed that the PD had knowledgeable facilitators (item 4) and provided opportunities to ask questions (item 5). Ratings related to collaboration, mix of activities, and modeling were also particularly strong. Evaluators observed that across schools and topics, PD facilitators modeled *Responsive Classroom* activities, showing participants what the strategies look and sound like through direct involvement. They also periodically encouraged participants to put on their “teacher hat” to discuss how the strategies could be used. On the other hand, 24.2% of respondents disagreed/strongly disagreed that the activities were engaging and 22.5% disagreed/strongly disagreed that it was worth their time.

We conducted several analyses to see whether different groups of educators responded differently to *Responsive Classroom* PD. First, we examined school role, separating respondents into special educators, regular educators, or educators who worked with full schools (i.e. as special teachers or administrators). We found a significant relationship between the respondents’ role in the school and their satisfaction with the training. Special education teachers were more likely to disagree (41.4%) that

the training was worth their time compared to regular education teachers (18.4%) or schoolwide teachers/administrators (9.1%) They were also more likely than others to disagree that the training used relevant materials (29.6% vs. 12.8% for regular educators and 7.3% for schoolwide educators).¹⁰ We also examined whether participants from different grade levels or different levels of background with *Responsive Classroom* responded differently. No significant differences were found.

Educators attending the summer institutes chose to sign up, sometimes with encouragement from supervisors. On the other hand, school-year PD activities in the partner schools were mandatory. We disaggregated our data by session type (summer vs. school year). Participants in summer sessions reported more engagement and higher levels of overall satisfaction. Some cells were too small to permit statistical tests, but the differences are clear. For example, 100% of summer participants agreed that the activities were engaging, compared to 65.9% of school-year participants. Likewise, 100% of summer participants said the workshop was worth their time, compared to 68.0% of school-year participants. The nature of attendance (voluntary vs. mandatory) may contribute to these differences.

Participants were asked to describe their favorite parts of the *Responsive Classroom* PD as well as any they would recommend changing.

Table A42
Participant Favorite Parts of the Responsive Classroom Workshop

Response (Category)	Number	Percentage
Collaboration with colleagues/classmates	88	33.1
Application of PD to the classroom	82	30.8
Engaging activities	54	20.3
Modeling of RC strategies	36	13.5
Facilitator	35	13.2
Other	32	12.0
RC philosophy/everything	12	4.5
Book used	9	3.4
Negative comment	6	2.5

Note: Percentages indicate the percent of respondents (N=233) providing this response.

Participating in *Responsive Classroom* training could foster community, both within schools and in institute groups. For instance, one teacher *“enjoyed the interaction with colleagues who I don’t get to see enough of during the course of the school day”* and a summer participant noticed *“the strong sense of community and belonging that developed in this classroom.”* Participants also felt the strategies could easily be applied: *“everything we did was productive and useful in the classroom.”* Other comments

¹⁰ Chi-square test indicates a statistically significant relationship between school role and training worth my time, $\chi^2(2, N=272) = 21.5, p=.000$. Chi-square test indicates a statistically significant relationship between school role and training used relevant materials, $\chi^2(2, N=275) = 14.0, p=.001$.

focused on engagement, modeling of activities, and facilitation, and these themes connected with each other. One participant felt the workshops *"kept me engaged and wondering about the next activity"*; another called out the *"mix of student-hat vs. teacher-hat learning"*; and several praised the facilitator as *"just amazing"* and *"very thoughtful in how she plans activities for us...I find myself thinking about things she has taught me when I am with my students."* Finally, some teachers talked about the whole philosophy of *Responsive Classroom* as their "favorite part"; this theme came through in interviews as well. Some teachers felt affirmed by the philosophy and others, productively, challenged: *"it pushed me outside my comfort zone. It made me think about 'how?' instead of 'there's no way my kids...'"*

Table A43
Participants' Suggestions for Changes to the Responsive Classroom Workshop

Response (Category)	Number	Percentage
Nothing, not sure, or only positive comments	78	31.2
Change logistical or design issues with PD (e.g., pace, grouping, physical environment, order of activities)	78	31.2
Make PD more engaging	45	18.0
Make activities specific to school role/context or differentiate	35	14.0
Make it longer or offer more sessions	25	10.0
Use more video or live examples	12	4.8
Improve facilitation	11	4.4
Provide more resources	7	2.8
Address strategies for discipline	6	2.4

Note: Percentages indicate the percent of respondents (N=250) providing this response.

Some of these themes are the inverse of the previous table, i.e. some participants were not engaged, did not find the activities relevant, or did not praise the facilitator. Suggestions for improvement were much more robust and common from the school-year surveys vs. the summer institutes. Comments related to logistics/design and engagement included requests for *"more time to practice and plan how to implement"* and *"more interaction among groups."* One participant commented: *"As teachers, we are constantly moving and engaging others. To sit for 3.5 hours with no engaging activities was very difficult."* Evaluator observations documented that summer participants were consistently highly engaged but in school sessions this varied depending on the time, composition of the group, dynamics with the facilitator, and other factors. One activity was reading from *Responsive Classroom* books, and participants tended to critique this (*"I am a good self learner and can read on my own time"*). On the other hand they appreciated videos and wanted more of them. Finally some comments reflect teacher doubts about how *Responsive Classroom* would work for their students; they wanted greater attention to issues of age (i.e., middle school), cultural diversity, special education, or discipline issues.

Survey comments indicate the most and least effective aspects of *Responsive Classroom* PD from participants' perspectives. Administrators also shared their thoughts on this question. In general, two principals were highly positive about their school's experience with *Responsive Classroom*, one was

mixed, and two were fairly critical, although they also acknowledged benefits. Opinions on the PD varied, with some effusively praising it and others wishing it were more engaging or reflective of their school needs or demographics. In direct and indirect comments, administrators generally expressed the belief that immersion in *Responsive Classroom* (i.e. through summer institutes) works best. Finally, they tended to express more satisfaction about the coaching component compared to the group PD.

Indeed, comments from coaching participants were unanimously positive and reflected satisfaction with the experience (even though one interviewee said s/he did not enjoy the group training). All participants reported feeling comfortable interacting with the coach, describing her as “very positive,” “helpful” and “respectful.” Furthermore they found the coach’s feedback beneficial, practical, and accessible. The coach was available after individual sessions by email.

Participant learning. Data to address this question comes from participant surveys, coaching interviews, and administrator interviews. We present the surveys first, then turn to the interviews.

Table A44

Participants’ Ratings of the Responsive Classroom Learning Outcomes

After participating in the PD activity:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
11. I learned new strategies for working with students. (N=290)	130 (44.8%)	118 (40.7%)	36 (12.4%)	6 (2.1%)	3.28 (.76)
12. The strategies taught in this PD are applicable for my students. (N=288)	139 (48.3%)	112 (38.9%)	31 (10.8%)	6 (2.1%)	3.33 (.75)
13. The PD showed me practical examples of how to use the strategies. (N=291)	139 (47.8%)	119 (40.9%)	28 (9.6%)	5 (1.7%)	3.35 (.72)
14. I know where to find resources/materials to support Responsive Classroom. (N=290)	107 (36.9%)	154 (53.1%)	25 (8.6%)	4 (1.4%)	3.26 (.69)
15. I feel confident in my ability to implement new strategies from the PD. (N=289)	80 (27.7%)	175 (60.6%)	31 (10.7%)	3 (1.0%)	3.15 (.64)
16. I will have to pick and choose which strategies from the PD to use. (N=285)	47 (16.5%)	188 (66.0%)	44 (15.4%)	6 (2.1%)	2.97 (.64)
17. Responsive Classroom fits well with other initiatives in my school. (N=287)	81 (28.2%)	168 (58.5%)	28 (9.8%)	10 (3.5%)	3.12 (.71)
18. Lack of resources at my school (e.g., time, materials, support) will make the strategies we covered difficult to implement. (N=288)	16 (5.6%)	81 (28.1%)	151 (52.4%)	40 (13.9%)	2.25 (.76)
19. I would like to receive coaching on Responsive Classroom. (N=277)	16 (5.6%)	106 (38.3%)	91 (32.9%)	33 (11.9%)	2.60 (.91)

For six of these items, 85% or more of participants gave a positive rating (agree/strongly agree). These include statements related to learning new strategies (item 11), finding the strategies applicable (item 12), seeing practical examples (item 13), knowing where to find additional Responsive Classroom resources (item 14) and feeling confident about using the strategies (item 15). In addition, 86.7% of respondents agreed/strongly agreed that Responsive Classroom fit well with other initiatives in their schools. Participants were asked to identify these other social curriculum initiatives. The most common was Positive Behavior Supports (PBS), which 70.5% of participants stated was used in their school.

We also conducted disaggregated analyses for some of these items to see if different groups of participants responded differently. Again, special educators voiced more concern. They were less likely to agree that the strategies taught were applicable for their students.¹¹ Also similar to the satisfaction data, respondents from the summer sessions generally responded more positively to these items than those from school-year sessions. Prior level of *Responsive Classroom* training did not matter (i.e., those with a lot, a little, and no prior experience were equally likely to report learning new strategies). Finally, educators who reported that PBS was active in their schools were *more likely* to feel that Responsive Classroom fit well. They were also more confident about implementation.¹²

Table A45

Participants' Expectations for Student Responses to Responsive Classroom Strategies

Response (Category)	Number	Percentage
Students will enjoy or respond positively	153	57.5
Students will need time to adjust	41	15.4
Students' participation or engagement will improve	37	13.9
Participant did not know or offered mixed/neutral response	31	11.7
Students will interact better with peers	22	8.3
Students will respond negatively	15	5.6
Students will behave better	10	3.8
Strategies are not appropriate for students	9	3.4

Note: Percentages indicate the percent of respondents (N=266) providing this response.

Most comments described anticipated or observed positive reactions. For instance, one teacher *"noticed a significant change in the overall tone of my classroom between last year and this year since receiving PD and coaching."* Others felt more specifically that *Responsive Classroom* could improve engagement, peer interactions, or behavior: *"I think they will be more ready to learn academics and be more self-regulated and self-motivated,"* and *"I have noticed students improve interactions with peers in*

¹¹ Chi-Square test indicates a relationship between school role and applicable strategies, χ^2 , (2, 272) = 9.9, $p=.007$

¹² Chi-Square test indicates a relationship between PBS use and RC fits with other initiatives, χ^2 , (2, 287) = 15.9, $p=.000$. Chi-Square test indicates a relationship between PBS use and confidence to implement, χ^2 , (2, 289) = 27.3, $p=.000$.

and out of the classroom.” On the other hand, some participants anticipated a difficult transition period: *“I think they may not be receptive at first but will grow into these strategies.”* Others expected mixed or negative responses: *“They may respond well but will probably ignore me or say ‘I don’t care.’”* Some comments express the view that *Responsive Classroom* would not work well for older students, students with disabilities, or students with behavioral challenges.

Table A46

Participants’ Expected Challenges to Implementing Responsive Classroom Strategies

Response (Category)	Number	Percentage
Challenges due to personal characteristics or skills	65	27.7
Challenges due to school context or role	61	26.0
Challenges due to student characteristics	54	23.0
Challenges due to time constraints	39	16.6
No challenges or unsure	28	11.9
Other	4	1.7

Note: Percentages indicate the percent of respondents (N=235) providing this response.

The largest group of comments relate to personal characteristics or skills. In a variety of ways, teachers described the challenge of shifting their own approach to the social curriculum: *“It is going to be a bigger adjustment for me rather than for the kids.”* Changing one’s use of language was seen as especially complex. Many also described contextual challenges in the school, ranging from their role (e.g., specials teacher) to the daily or annual calendar: *“there is no structure/time in the schedule for morning meeting”* and *“it is late in the year. I can’t wait to use this at the beginning of the school year.”* Some, especially from non-partnership schools, also anticipated challenges getting buy-in from colleagues or administrators: *“My administration doesn’t know much about RC and I am not sure they will be supportive of my efforts in implementing it.”*

Coaching participant and administrator interviews provide more insight into teacher development and instructional change. Both sets of interviews describe the power of the one-on-one, embedded coaching. As one administrator stated,

To have someone provide you with immediate feedback was very powerful and helped teachers maybe recognize behaviors they either didn’t realize or maybe they just weren’t aware of. And therefore when someone brings that to your attention – and also in a very non-threatening way... – I think that also helped with teacher [receptiveness].

In some cases, coaching helped open teachers’ minds to *Responsive Classroom*. The coach was able to suggest and show alternatives, such as this example about interactive modeling:

So when I was going through what I expected for them to do during their activity,

she showed me how even though it takes a lot of time to begin, it makes it quicker because I wouldn't have to go back and recheck and redo and fix up.

In other cases, the coaching augmented or affirmed strategies teachers were already attempting. One interviewee was not sure her use of time outs *"was the Responsive way to do it."* After meeting with the coach, she felt, *"Okay, I am doing it right.... I am going to continue to do this."* One teacher commented, *"I guess it mostly gives me a name for what I thought [already]."*

Of course, the challenge this activity took on was to try to change entire schools, where some staff members would more naturally embrace *Responsive Classroom* than others. In some partnering schools, *Responsive Classroom* was already well established while it was new in others. Where *Responsive Classroom* was already present, administrators generally wanted to make its use more consistent and/or deeper. Some identified the strategies as a good fit for their students' needs:

If we change our approach, I think it would really help [students] and we could get so much further along in helping them with their academics. Because once they lose it, once you hit their trigger, they're done and they have a meltdown, the day is lost.

Two administrators emphasized the difference between *Responsive Classroom* and past practice or teacher beliefs about what was appropriate. They described specific practices that teachers felt they were being asked to abandon (e.g., holding students in from recess). Teacher resistance ensued. While these situations differed, in both cases the leaders felt *Responsive Classroom* was *"a huge shift"* and felt more could have been done, to show *"how it could really be helpful for the kids and for [the staff]."*

Ultimately, teacher implementation varied within and across schools. One principal identified across-the-board implementation and said his/her goals were met. However, s/he also noted that some teachers went beyond: *"the more time they invest in the PD, the greater the outcome, the greater the shift."* This idea was repeated in other interviews, and principals in those schools further described some educators who only used *Responsive Classroom* for compliance: some *"really dug into [it]...whereas others I think just walked away and sort of were skimming the surface."* Administrators perceived that implementation was more successful when teachers had prior background or taught lower grades. In one school, the administrator expressed hope for the future but acknowledged, *"we have made great strides but we are not yet at the tipping point"* where most educators had 'bought in.' Finally, one administrator stated there was as yet little implementation, except at one grade level.

Student achievement. In interviews teachers and administrators were asked about student responses. Teachers tended to describe positive reactions though in some cases transition time was needed. Comparing the 2014-15 school year to prior years, one teacher noted, *"I feel like we really did create a community, whereas before it wasn't."* Another saw links between climate and academic changes: *"I've seen [students] mature a lot more. I've seen them really take ownership of behaviors and choices and academics because I took the conversation off my shoulders and put it onto their shoulders."*

Administrators, too, recounted positive impressions of student response, especially where teachers were implementing consistently, but felt it was too soon to see solid impacts. Several articulated that this was a process and change would not occur overnight. Only one made a statement like this: *"I have definitely seen some positive growth in the area of interacting appropriately with each other, and overall a decrease in our behavior-like referrals."*

Finally, we collected quantitative data related to student outcomes and school climate. Schools partnering on this grant were asked to set targets so that data from before and after the grant year could be compared. They used school climate, discipline, or teacher evaluation data. Of course, schools were involved in many concurrent PD or school improvement efforts. No claims about causation are made. Five schools provided school discipline or school climate data for the 2014-15 school year, that could be compared to the prior year. Two schools responded to the request but were unable to provide comparable data. Of the data received, four included quantitative data on out-of-classroom referrals as well as suspensions. These are shown in Table A47.

Table A47

School Disciplinary Action Data for 2013-2014 School Year and 2014-2015 School Year

School	2013-2014					2014-2015				
	Total number and (%) of students receiving out-of-class referrals			Total number out of class referrals	Total number suspensions	Total number and (%) of students receiving out-of-class referrals			Total number out of class referrals	Total number suspensions
	0-1	2-5	More than 6			0-1	2-5	More than 6		
New Castle Elementary	455 (84.1%)	56 (10.4%)	30 (5.6%)	541	125	444 (80.0%)	67 (12.1%)	44 (7.9%)	593	149
Pleasantville Elementary	447 (91.8%)	26 (5.3%)	14 (2.9%)	513	48	426 (97.0%)	12 (3.0%)	2 (0.6%)	459	26
Southern Elementary	810 (91.9%)	48 (6.5%)	23 (2.6%)	1041	88	775 (89.1%)	59 (6.8%)	36 (4.1%)	1405	156
Warner Elementary	434 (74.3%)	26 (11.5%)	83 (14.2)	1653	238	426 (72.9%)	62 (10.6%)	96 (16.4)	1770	165

Pleasantville Elementary showed decreases in the total number of referrals and suspensions between 2013-14 and 2014-15. There were also decreases in the number and percentage of students at the top

of the behavior pyramid (i.e., receiving 2-5 or 6+ referrals) and an increase in those receiving 0-1 referrals. This is a positive change. At Warner, results were mixed. The number of suspensions decreased but the total referrals increased, and more students received 6+ referrals in 2014-5 compared to the previous year. At both New Castle and Southern, the total number of suspensions and referrals increased from 2013-14 to 2014-15 and the distribution of students shifted upwards on the 'pyramid,' with more students receiving a higher number of suspensions than the previous year.

ASPIRA provided data related to instructional practice, since the school's stated goal was to improve student engagement and instructional time. The data source was DPAS II ratings and narratives but these were described qualitatively not quantitatively. At the end of the 2014-15 school year, a school administrator submitted this statement:

The qualitative DPAS II data for the 2014-15 school year indicates that modest gains were made in the targeted areas (time on task, engagement during student-centered projects and learning activities, and effective and smooth transitions). Our educators need additional training and time to implement the full range of Responsive Classroom practices. We anticipate that this will lead to more consistent gains school-wide.

Educator Design Workshop

Introduction. UD engineering, computer science, math, and education faculty collaborated to develop and lead the Educator Design Workshop. This activity aimed to immerse K-12 educators in hands-on discovery of STEM and in so doing to build participants' technical knowledge and confidence as designers and makers. It also encouraged educators to integrate design concepts into their instruction.

This activity had two parts. In February, March, and May, half-day workshops were held on three STEM topics: computer programming, electronic textiles, and 3-D printing and prototyping. After a brief introduction to the topic, participants spent most of the time engaged in hands-on discovery projects. For instance in February, they programmed computer games using Scratch; in March, they designed and created a stuffed animal that could light up. There were also opportunities for participants to discuss how to use workshop activities or concepts in their classrooms. The workshops took place at UD and incorporated many university resources such as the technology-rich Design Lab, as well as support from UD STEM students, who attended workshops and helped teachers with the projects.

In the summer, participants returned for a more extended workshop, which occurred over three consecutive full days. Participants were introduced to the four phases of the engineering design cycle and to a variety of case studies about how it could be incorporated into courses and curricula. The emphasis was on "design of design" – how to effectively develop, teach, and assess an engineering design challenge with students. Interwoven with the content and pedagogical discussions, participants also developed a project. They formed groups and were encouraged to engineer a response to a "wicked problem" in healthcare or sustainability. Through this abbreviated hands-on design challenge, participants gained exposure to the different phases of the process.

Space in the workshop was limited, and ten teachers from high-need elementary and middle schools enrolled. One teacher from a non-high need school also participated in some sessions, with no stipend. Participants were asked to commit to the whole series (i.e., three Saturdays and a three-day summer workshop) although in practice attendance varied slightly. One educator who attended the Saturdays did not return in the summer. Surveys were administered after each Saturday session, to permit formative adjustments, and a longer survey was administered at the end of the summer workshop. Items related to satisfaction remained the same across all surveys, so we present aggregate results here. For items related to instructional application, we focus on data collected at the end of the series.¹³ One limitation is that only seven of the ten participants were present to take the final survey.

Participant satisfaction and professional development design. Table 48 shows the aggregated results for items related to satisfaction and professional development design across four sets of surveys.

Table A48

Participants' Ratings of Educator Design Workshop Design (Aggregated Across Four Workshops)

The workshop:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
1. Was engaging and kept my attention.	31 (88.6%)	4 (11.4%)	0 (0.0%)	0 (0.0%)	3.89 (0.32)
2. Included a good mix of activities.	31 (88.6%)	4 (11.4%)	0 (0.0%)	0 (0.0%)	3.89 (0.32)
3. Used materials that are relevant to my teaching situation.	25 (71.4%)	9 (25.7%)	1 (2.9%)	0 (0.0%)	3.69 (0.52)
4. Had knowledgeable facilitators/instructors.	29 (82.9%)	6 (17.1%)	0 (0.0%)	0 (0.0%)	3.83 (0.38)
5. Gave me sufficient opportunities to ask questions.	34 (97.1%)	1 (2.9%)	0 (0.0%)	0 (0.0%)	3.97 (0.17)
6. Encouraged meaningful collaboration with colleagues.	32 (91.4%)	3 (8.6%)	0 (0.0%)	0 (0.0%)	3.91 (0.28)
7. Tried to cover too much in the time available.	6 (17.1%)	6 (17.1%)	14 (40.0%)	9 (25.7%)	2.26 (1.02)
8. Modeled effective instruction.	22 (62.9%)	10 (28.6%)	3 (8.6%)	0 (0.0%)	3.54 (0.65)
9. Took place in a space that met our needs.	32 (91.4%)	3 (8.6%)	0 (0.0%)	0 (0.0%)	3.91 (0.28)
10. Addressed a topic that was new to me. ^a	23 (67.6%)	7 (22.7%)	4 (27.3%)	0 (0.0%)	3.56 (0.69)
11. Was worth my time. ^a	32 (94.1%)	2 (5.9%)	0 (0.0%)	0 (0.0%)	3.94 (0.24)

N=35. This is a composite of four sets of surveys.

^a N=34. One participant selected between "agree" and "disagree" for item 10. One participant skipped item 11.

¹³ Summaries of the February, March and May surveys were shared with project staff soon after each event.

These ratings are overwhelmingly positive and show persistently high levels of educator satisfaction across the four workshops. Participants found the workshops especially engaging (item 1) and appreciated the activities and the spaces used (items 2 and 9). Teachers were observed to be highly engaged throughout the workshop and involved in a number of novel and creative hands-on activities.

Participants strongly agreed they had opportunities for collaboration and questioning (items 5 and 6). Here, the small group size and level of faculty/student support helped. Overall the workshop environment was observed to be relaxed and conducive to questions. The only area of any concern in these ratings is the workshops' scope. Over a third of respondents felt the workshops tried to cover too much in the time available. The Saturday workshops crammed a great deal of information and practice into just four hours each, and the pace in the summer was also very quick.

Participants were asked to identify their favorite part of the Educator Design Workshop:

Table A49

Participant Favorite Parts of the Educator Design Workshop (Aggregated Across Four Workshops)

Response (Category)	Number	Percentage
Hands-on opportunities to try skills and use technology	16	45.7
Facilitators, students, or general support available	10	28.6
Content/information or resources shared	10	28.6
Collaboration with peers	6	17.1
Applicability of activities to classroom	6	17.1
Organization/environment of workshop	5	14.3
Inspiration for teachers or students	3	8.6
Other or everything	2	5.7

Note: Percentages indicate the percent of respondents (N=35) providing this response, in the aggregated survey set.

By far the largest group of responses pertained to the hands-on nature of the workshops. Referencing Scratch, a respondent stated, *"I loved having time to try it out myself"* and another enjoyed *"working on a real project."* Another consistent theme across the data was the level of support provided: *"Small # of participants. Amount of helpers available."* On a summer survey, one person wrote, *"the best part was the faculty – very knowledgeable and enthusiastic...very helpful, and all had strong knowledge base."* Other comments focused on the content of the workshops or the collaboration with peers. A variety of collaborative activities were used, such as pair programming or peer feedback on engineering designs. Finally, several participants wrote positive comments about how the workshop could inspire both teachers and students:

What I liked best was it took me out of my comfort zone and allowed me to grow in knowledge about subject areas I wouldn't have covered otherwise.

It is important to emphasize to our students that our activities lead to what they do in college and careers.

Workshop observations document a persistent theme: to expand access to STEM, we need to make activities relevant to students' lives and empower them to delve into it.

Across all four sets of surveys, participants had only minimal suggestions for improvements to the Educator Design Workshops (Table A50).

Table A50
Participant Suggested Changes for the Educator Design Workshop (Aggregated Across Four Workshops)

Response (Category)	Number	Percentage
Not sure, none or only positive comments	12	40.0
Allocate more time	10	33.3
Adjust the activity/project/presentation	5	16.7
Enhance materials/examples/resources	4	13.3
Enhance technology	2	6.7
Emphasize pedagogical application	1	3.3

Note: Percentages indicate the percent of respondents (N=30) providing this response, in the aggregated survey set. Five respondents left this item blank.

Consistent with the quantitative data, a third of comments suggest increasing the length of workshops. A smaller group suggested revisions to workshop activities such as, "different make projects with the same materials" "could we please see actual printing happen?" or "more time to complete end of project presentations." In the summer session, the final presentations were observed to be rushed and somewhat disjointed. Final reflection activities on the Saturday sessions were often cut short. Some participants suggested improvements to workshop materials or technology but these were minor, such as requesting wifi passwords ahead of time or blank copies of handouts.

Participant learning and instructional planning. Here we focus on data from the surveys administered at the end of the summer workshop, which explored these issues in more depth. Participants' ratings of their learning outcomes from the *summer* workshop are shown in Table A51.

Table A51
Participant Ratings of Learning Outcomes for the Summer Educator Design Workshop (Aggregated)

After participating in the PD activity:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
12. I learned new approaches for teaching students engineering design.	6 (85.7%)	1 (14.3%)	0 (0.0%)	0 (0.0%)	3.86 (.38)
13. The case studies helped me understand engineering design better.	6 (85.7%)	1 (14.3%)	0 (0.0%)	0 (0.0%)	3.86 (.38)
14. I see the connections between the different topics we covered this year.	7 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	4.00 (.00)

After participating in the PD activity:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
15. I will be able to use the material covered in the workshop with my students.	5 (71.4%)	2 (28.6%)	0 (0.0%)	0 (0.0%)	3.71 (.49)
16. I have already started to implement some of what I learned in the PD this spring with my students.*	2 (33.3%)	3 (50.0%)	1 (16.7%)	0 (0.0%)	3.17 (.75)
17. Lack of resources at my school will make the techniques we covered difficult to implement.	1 (14.3%)	4 (57.1%)	2 (28.6%)	0 (0.0%)	2.86 (.69)
18. I am confident in my ability to implement ideas and practices from the workshop in my classroom next year.	4 (57.1%)	3 (42.9%)	0 (0.0%)	0 (0.0%)	3.57 (.53)
19. My design challenge still needs much more development.	1 (14.3%)	5 (71.4%)	1 (14.3%)	0 (0.0%)	3.00 (.58)

N=7.

* One participant selected in between "agree" and "disagree." This rating was removed from analysis.

Respondents overwhelmingly agreed that they learned new approaches to teaching engineering design and that the case studies were helpful (items 12 & 13). Presenting multiple case studies instead of one extended example (as was done in 2014) gave educators more ideas for application and highlighted opportunities in different disciplines. However, it also meant sacrificing some depth. Six participants agreed that their design challenge needed much more development (item 19). Combining content, pedagogy, and hands-on experience in a single workshop was ambitious. It was not expected that teachers would complete a full prototype; instead, the purpose was to give them a taste of the experience. As one presenter put it, *"when you go back to your school, you'll have something to show the kids. I made this. You can too."* Nevertheless, presentations made clear that educators still had much work to do with their designs and their instructional plans.

Respondents gave the highest possible ratings to the coherence of the workshop series (item 14). Facilitators were observed drawing connections between the skills and technologies taught (e.g., between Scratch and Arduino). Yet while educators reported recognizing the connections, there was limited evidence of transfer in their responses to the final task. Participants were asked how, if at all, the winter/spring workshops influenced their thinking about the summer design challenge; responses are in Table A52, below.

Table A52

Influence of Prior Experiences on Participants in the Summer Design Workshop

Response	Number	Percentage
Gained motivation for the summer design challenge	3	42.9
Taught or will teach ideas from spring to students	2	28.6
Did not use spring skills/concepts in summer design challenge	2	28.6
Used specific spring skill in summer design challenge	1	14.3

N=7.

One participant used 3D printing to make a game piece. Others described much more general transfer, such as: *"I was aware of more possibilities that I could access in my design process. I appreciate the time, discussions and hands-on activities that gave me the experience which prepared me for the 3-day workshop."* However, one participant identified barriers to using technologies s/he had learned: *"they might be ways to solve our challenge but we don't have supportive resources at our school to ensure we could actually use them in the challenge we designed."* Indeed, the majority of respondents identified school resources as a limitation (item 17 Table A51). These include materials, technology, space, and also human resources. In informal conversation, one participant described how it would be challenging to teach this material or do these activities with a large class, saying *"look at the bodies it takes to teach it to adults."* As a corollary to this statement, this resource-rich PD activity could only reach a small number of educators.

All respondents felt they would be able to use workshop material with their students (item 15, Table A51) and indeed some had already done so (item 16, Table A51). We asked teachers to describe how they planned to teach the design challenge and integrate it with existing curricula (Table A53).

Table A53

Participants' Plans to Integrate Engineering Design Challenge into their Classes or Curriculum

Response	Number	Percentage
Integrate into a specific science unit	3	42.9
Adapt or differentiate process for students	3	42.9
Integrate challenge into a social studies unit	1	14.3
General statement about plans to integrate	1	14.3

N=7.

Several responses were quite specific. Since educators were able to develop their own challenges, it is logical that they saw many opportunities for integration into science or other units. Other responses focused on ways the challenge could be adapted or lessons could be drawn from it more broadly:

I plan to implement this in my classroom much like it was presented here but scaled down a notch or two, for elementary

I would like to share this project with kids and let them know that we make mistakes and learn to re-evaluate

Throughout the workshop, participants and facilitators discussed the need to provide different levels of scaffolding and differentiation, depending on the age and abilities of the students.

In general, participants reported that they expected students to respond positively to engineering design (Table A54).

Table A54

Participants' Expectations about Student Reactions to the Educator Design Workshop Strategies

Response	Number	Percentage
Students would respond positively or be engaged	4	57.1
Students would enjoy the hands-on aspect	2	28.6
Students would enjoy the novelty	2	28.6
Students would lose focus during some stages	1	14.3

N=7.

Most of the comments anticipated positive responses: *"I think my students would respond positively because it will be engaging and challenging. It will be different from what they're used to."* Only one caution was heard: *"some students might get unfocused during the stage in the design process that requires them to do background research or define the metrics – I think middle school kids might be challenged with staying focused on the end goal. It is my job to monitor, observe, and scaffold."*

Finally, educators were asked what challenges they anticipated in using what they learned.

Table A55

Participants' Anticipated Challenges to Applying Educator Design Workshop Strategies

Response	Number	Percentage
Constraints of time or existing curriculum	3	42.9
Funding or materials needed	3	42.9
Time required for preparation/planning	2	28.6
Other contextual challenges	2	28.6

Percent is the number of respondents (N=7) providing this response. One respondent gave multiple responses.

A group of responses cited time limits or competing demands, especially since participants had experienced how much time it takes to do an engineering design challenge well: *"Main challenge is to stay within a timeframe that ensures success in the design process but also does not put teachers way behind in the content pacing of the unit."* Observation notes indicate that concerns about time, alignment and the freedom to be creative (or lack thereof) were repeatedly brought up. In the first workshop, one participant commented: *"This is wonderful but in the classroom, we can't just stop and do things...teachers have to be accountable for what they are teaching."* When such concerns were raised, facilitators and fellow participants engaged them and attempted to provide practical ideas. Information about the alignment between activities and NGSS was also shared.

The cost of materials was also seen as a barrier. During the workshop, resources were bountiful. Materials for e-textiles cost about \$50 per participant, and in the summer one facilitator went on an evening shopping trip to procure any specific materials teachers requested for their project. Much as teachers might wish, this service is not available during the year. Other challenges cited included *"lack of enthusiasm from colleagues"* and lack of instructional and storage space.

Instructional change and student achievement. Much of this activity occurred during the summer, leaving insufficient time to track the impact on teachers' instruction and student achievement.

Delaware Alternative Routes for Educators (DARE)

Introduction. The Delaware Alternative Routes for Educators (DARE) program is an alternative route to certification for teachers of world languages or special education working in high-need middle or high schools. DARE aims to recruit, select, train, develop and retain effective teachers. The program includes a combination of supports: an intensive summer institute, ongoing professional development, one-on-one coaching and mentoring, and tuition support for UD coursework. The first cohort of DARE teachers was recruited in summer 2014, under a previous grant. The current grant supported activities for this first cohort during the 2014-15 school year, as well as the recruitment, admissions, and initial training of a second cohort of teachers.

One challenge of evaluating DARE is its close relationship to two other programs. The first is UD's regular Alternative Routes to Teacher Certification (ARTC) which prepares secondary teachers in critical need areas in schools regardless of high need status. The second is the Delaware Transition to Teaching Partnership (DT₃P), a United States Department of Education funded program that is similar to DARE but is limited to a more stringent federal eligibility for "high need LEAs." DARE and DT₃P are run by the same individuals at UD and include the same admissions process. In practice, candidates are accepted and then depending on where they obtain positions (subject and school eligibility) their program placement is finalized. A related challenge is the very small sample size. As discussed below, the first DARE cohort ultimately only included two candidates.

DARE is somewhat different from the other Title II activities. We first report information about candidate recruitment, admissions and placement before examining candidate satisfaction.

Recruitment, admissions, and placement. Cohort 1 recruitment and admissions occurred before the start of the current Title II grant. By October 2014, there were two DARE candidates who had obtained teaching positions in eligible schools, both as special educators. One taught high school ELA at Lake Forest High School. The second taught middle school social studies at Family Foundations Academy; midyear, s/he transitioned to a more specialized special education role. Both candidates retained their positions for the 2015-16 school year.

Three educators who had been accepted into DARE Cohort 1 completed the summer institute but did not obtain teaching positions for 2014-15. They included two world language and one special education teacher. After the summer institute (i.e., in October 2014) they were transferred to the ARTC program. A sixth candidate was accepted into DARE but did not complete the summer institute.

Reporting for DARE Cohort 1 will be limited due to the small number of candidates (2) and the need to protect their confidentiality.

Recruitment and admission for DARE Cohort 2 began in winter 2015 and was coordinated with DT₃P. As discussed above, some candidates were admitted who were eligible and willing to join either program, depending on the position they obtained. These are considered initial members of the DARE cohort. DARE established several criteria for candidates, outlined in Table A56 below. The number and percent of candidates meeting each criterion is also indicated. Data came from project staff.

Table A56

Initial Members of DARE Cohort 2 Measured Against Admissions Criteria

Criterion	Number Meeting	Percent Meeting
Undergraduate degree	10	100.0
GPA of at least 3.0	9	90.0
Passed PRAXIS CORE or exempt from this requirement due to other test scores	9	90.0

N=10.

From these ten admitted candidates, four obtained teaching positions that qualified them for DARE. They were placed at the Delaware Met Charter School (1), Dover High School (1) and William Penn High School (2). Three taught core content special education and one taught world languages (Spanish). All four of these candidates successfully completed the DARE summer institute.

The other six members of the original cohort had various outcomes. One candidate had not yet obtained a position by September 30, 2015 and was on a leave from the program due to personal circumstances. Two candidates who had been willing to join DARE ultimately accepted science teaching positions, which meant they both ended up as DT₃P candidates instead. Finally, three candidates did not complete the summer institute. One of these never showed up, one attended just one day of training, and one formally withdrew in August to pursue a different career.

Participant satisfaction. Participant satisfaction surveys were administered regularly to DARE participants. Cohort 1 candidates completed surveys after each of four workshops during the 2014-15 school year. These were half-day professional development workshops offered jointly with DT₃P. All four focused on aspects of teaching in which candidates needed extra support, based on observations and feedback from all stakeholders. Topics included classroom management (November), lesson planning (December), developing student surveys (February) and using survey results to improve classroom environment and engagement (March).

Due to the small number of DARE candidates, we reported aggregated results for the survey items that remained the same across these four instruments (Table A57)

Table A57

Cohort 1 DARE Candidates' Ratings of Workshops (Aggregated Across Four Workshop Surveys)

The workshop:	Strongly Agree	Agree	Disagree	Strongly Disagree	Mean (SD)
a. Was engaging and kept my attention.	8 (66.6%)	4 (33.3%)	0 (0.0%)	0 (0.0%)	3.67 (0.47)
b. Included a good mix of activities. *	6 (54.5%)	5 (45.5%)	0 (0.0%)	0 (0.0%)	3.55 (0.50)
c. Covered content that I will be able to use in a classroom.	8 (66.6%)	4 (33.3%)	0 (0.0%)	0 (0.0%)	3.67 (0.47)
d. Covered too much in too short a time period.	0 (0.0%)	2 (16.7%)	4 (33.3%)	6 (50.0%)	1.67 (0.75)
e. Provided me with opportunities to collaborate with fellow candidates and DARE staff.	8 (66.6%)	4 (33.3%)	0 (0.0%)	0 (0.0%)	3.67 (0.47)
f. Was redundant with other professional development I have received.	0 (0%)	2 (16.7%)	4 (33.3%)	6 (50.0%)	1.67 (0.75)
g. Had a clear purpose.	10 (66.7%)	2 (33.3%)	0 (0.0%)	0 (0.0%)	3.83 (0.37)
h. Provided sufficient opportunities for questions.	10 (66.7%)	2 (33.3%)	0 (0.0%)	0 (0.0%)	3.83 (0.37)
i. Was a good use of my time.	7 (58.3%)	5 (41.7%)	0 (0.0%)	0 (0.0%)	3.58 (0.49)

N=12.

* N=11. One participant skipped this question.

These ratings suggest high levels of candidate satisfaction and workshop relevance. In particular, candidates found that the workshops were engaging, had a clear purpose, and fostered both collaboration and questioning. A minority felt that the workshops attempted to cover more than was reasonable in four hours, or on the other hand that the topics were redundant with other PD.

Table A58

DARE Cohort 1 Most Useful Parts of Workshops (Aggregated Across Four Workshop Surveys)

Response (Category)	Number	Percentage
Process of creating tangible products	5	41.6
Collaboration with colleagues	4	33.3
Opportunity for reflection and examination of data	3	25.0
Support from coach or staff	2	16.7

N=12.

Three of the workshops were focused on creating tangible products (behavior plans, lesson plans, or student surveys). Candidates appreciated the time to think and develop a product and to walk away

with something they could use. Second, they appreciated collaborating with members of the cohort (e.g., through a *“critical friends”* protocol). DARE coaches and staff members also attended workshops and provided layers of support. One candidate found *“working through ideas with a coach”* most useful. Finally three comments referred to the value of reflecting on data and on individual performance. Across these comments, DARE workshops appear to provide time, space, and support for candidates to plan and reflect in ways that are not always possible during the hectic first year in the classroom.

Candidates were also asked for recommendations to improve the workshops (Table A59).

Table A59

DARE Cohort 1 Recommendations for Improving Workshops (Aggregated Across Four Workshop Surveys)

Response (Category)	Number	Percentage
No suggestions	4	36.4
Allocate more time (in general or for specific activities)	3	27.3
Address content area, grade level, or special education needs	2	18.2
Increase support for individual needs	1	9.1

N=11. One candidate skipped this item on one survey.

Relatively few substantive suggestions were offered. Reflecting the ratings (above), candidates expressed the need for more time. Others suggested grouping candidates by content area or *“address[ing] differing grade levels/learning environments”* so that the conversation could be more focused on particular needs. Extending this idea, one comment recommended *“allow[ing] really difficult cases one on one support”* so that they would not dominate group conversations or activities. Throughout the year, balancing the needs of the group and of individual candidates who were teaching in varied, often extremely demanding, environments was a challenge.

In addition to the workshops, candidates participated in UD courses during each semester and the summer session. Due to the small number of respondents, satisfaction data for these courses is not reported. Candidates were also observed and videoed by their coaches and then debriefed. Again, data about candidate satisfaction with coaching was obtained through interviews but will not be reported.

Participant learning, instructional change, and student achievement. Data related to these evaluation questions were collected through interviews with DARE candidates and their coaches at the end of the first year of teaching. We also collected survey data from school administrators and school/district mentors, as well as candidate DPAS II scores. Because only there were only two members of DARE Cohort 1 teaching, these data cannot be reported here.

Discussion

This section synthesizes across the results of the seven Title II activities, presented above, and considers overall satisfaction, professional development utility and quality, participant learning, and where data are available, instructional change and student learning. It then provides recommendations for the future.

In general, participants were satisfied with their Title II experiences. Across all activities, 82.8% of participants agreed/strongly agreed that the PD was a good use of their time. Nearly half strongly agreed. We see some satisfaction differences between and within activities; the most notable are differences between sessions of *Responsive Classroom*. This activity was fundamentally different from the others in that it required school wide participation – a much more challenging context for PD.

Effective PD is collaborative and engaging, involving educators in active learning and modeling strategies they can use in their work. All seven activities emphasized collaboration and 91.3% of all ratings related to collaboration were positive (agree/strongly agree). Overall, 81.4% of respondents agreed/strongly agreed that the activities were “engaging and kept their attention.” While this is a reasonable overall rate, some activities were more engaging than others. Engagement was lower where activities prioritized content delivery or moved at too slow a pace, and/or where educators were required to attend and sometimes resistant.

Once again, content focus was a highlight this year. Every activity except for *Responsive Classroom* had a disciplinary thrust, and *Responsive Classroom* emphasized the academic and learning benefits of a positive classroom culture. Several activities (e.g. Educator Design Workshop) emphasized cross-disciplinary connections. Alignment to the CCSS, NGSS, and Smarter Balanced was observed in the PD material and activities, if not always in participant work products.

Across all seven activities, we noted a variety of research-supported and creative designs, including:

- One-on-one, school-based coaching (Literacy Coaching at East Side Charter School, *Responsive Classroom*, DARE)
- Hands-on building and discovery (Educator Design Workshop)
- Use of technology to extend learning or increase access (Non-Fiction Writing Institute, Social Studies Webinars)
- Collective participation by teams of teachers from same schools (Literacy Coaching at East Side Charter School, *Responsive Classroom*, Writing Module Development, as well as some teams from Educator Design Workshop)

The variety of designs is striking. As documented through this report, each activity had strengths and areas it needed to improve. There were some design aspects on which the research is clear but the practices were uneven. For example, effective PD takes place over time. While most of the activities were sustained over months, a few used only a summer institute model (Non-Fiction Writing, some

Responsive Classroom institutes). Although planned as a series, the social studies activity in fact had most participants register for only one or two hour-long webinars.

Participants in Title II activities generally felt positive about what they learned; 87.2% agreed/strongly agreed they had learned new strategies.¹⁴ Pre and post surveys of participants in the module development suggest statistically significant changes in some areas of content and pedagogical knowledge. The PD activities modeled a range of instructional techniques that participants could observe, try out, and bring back to their classrooms; these were often cited as a highlight. All activities except for the social studies webinars provided at least some time and support for participants to plan for applying what they had learned, although the scope varied considerably. In some activities, participants submitted and/or presented plans that could be evaluated. These tended to be inconsistent, although they generally reflected effort and attempts at applying new knowledge. Our evaluation suggests that individual coaching was a particularly powerful lever for learning.

Some activities allowed us to systematically examine instructional change and others did not. This depended on the schedule and also the scope of the activity. Where activities were concentrated in a single school (Literacy Coaching, *Responsive Classroom*), administrator feedback provided an important window on implementation. Generally, we heard that teachers were attempting to apply new strategies and that those who invested more fully in the PD and/or coaching were having better results. Our data affirm that instructional change does not happen overnight and requires ongoing support from professional developers and from administrators.

Finally, student-level data were only collected for three activities. Student reading data at East Side Charter School were incomplete but the school met three of four targets for which data are available. For the writing module activity, student writing showed improvement at all three grade levels. School climate data in the *Responsive Classroom* schools were mixed, but we note that this measure was less tightly aligned to the PD than the others.

The 2015 Title II grant provided a wide range of activities, and we conclude by emphasizing that range. As described throughout this report, activities varied tremendously in scope as measured by the number of participants, the time span, and the total time allocation. They also used resources (e.g., facilitators, stipends) differently. Some engaged administrators; others did not. Some activities were open to all; others had limited spaces or application processes, and still others were required. The activities shared a goal to improve teaching and learning in high-need schools, but showed different theories of change about how to make this happen.

¹⁴ This item was not included on surveys for social studies webinars or DARE.

Recommendations

1. **Learn from what worked well in 2015.** As described throughout this report, every activity this year had some successes. We encourage program staff to identify and build not only on their own successes, but those in other parts of the grant as well. We see a major opportunity (missed in this past year) for program staff to regularly meet and share best practices.
2. **Increase coordination among parts of the project.** This year's Title II activities reached many educators across different settings, content areas, age levels, PD interests etc. Yet its breadth also meant a lack of cohesion. Sometimes, participants in a school were involved in multiple activities without understanding that they were part of the same grant. Staff members were not always aware of what others were doing. Again, these were missed opportunities for alignment, efficiency, and professional learning.
3. **Establish shared parameters for PD design.** We are *not* recommending that Title II PD be standardized; indeed, creative approaches to PD are a hallmark of this grant, and different participants are drawn to different activities. On the other hand, the wide design variations in the past year were not optimal. It is reasonable to require, for example, that all activities take place over time and/or require a certain level of follow-up with participants, or engagement with their administrators.
4. **Refine measures of instructional change and student achievement.** Program designers should be as specific as possible about the intended outcomes for participants and their students, and targets should be reasonable given the grant timeline.

References

Culham, R. (2014). *The Writing Thief: Using Mentor Texts to Teach the Craft of Writing*. Newark, DE: International Reading Association.

Guskey, T.R. (2012, October). *Making a Difference: Revolutionizing the Way We Plan and Evaluate Professional Development*. Pre-conference institute at the annual conference of the Association for Supervision and Curriculum Development. Los Angeles. CA.

Appendix 1: Standards

Learning Forward Standards for Professional Learning

Learning Communities:

Professional learning that increases educator effectiveness and results for all students occurs within learning communities committed to continuous improvement, collective responsibility, and goal alignment.

Leadership:

Professional learning that increases educator effectiveness and results for all students requires skillful leaders who develop capacity, advocate, and create support systems for professional learning.

Resources:

Professional learning that increases educator effectiveness and results for all students requires prioritizing, monitoring, and coordinating resources for educator learning.

Data:

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.

Learning Designs:

Professional learning that increases educator effectiveness and results for all students integrates theories, research, and models of human learning to achieve its intended outcomes.

Implementation:

Professional learning that increases educator effectiveness and results for all students applies research on change and sustains support for implementation of professional learning for long-term change.

Outcomes:

Professional learning that increases educator effectiveness and results for all students aligns its outcomes with educator performance and student curriculum standards.

Delaware Professional Teaching Standards

Learner Development

The teacher understands how learners grow and develop, recognizing that patterns of learning and development vary individually within and across the cognitive, linguistic, social, emotional, and physical areas, and designs and implements developmentally appropriate and challenging learning experiences.

Learning Differences

The teacher uses understanding of individual differences and diverse cultures and communities to ensure inclusive learning environments that enable each learner to meet high standards.

Learning Environments

The teacher works with others to create environments that support individual and collaborative learning, and that encourage positive social interaction, active engagement in learning, and self motivation.

Content Knowledge

The teacher understands the central concepts, tools of inquiry, and structures of the discipline(s) he or she teaches and creates learning experiences that make the discipline accessible and meaningful for learners to assure mastery of the content.

Application of Content

The teacher understands how to connect concepts and use differing perspectives to engage learners in critical thinking, creativity, and collaborative problem solving related to authentic local and global issues.

Assessment

The teacher understands and uses multiple methods of assessment to engage learners in their own growth, to monitor learner progress, and to guide the teacher's and learner's decision making.

Planning for Instruction

The teacher plans instruction that supports every student in meeting rigorous learning goals by drawing upon knowledge of content areas, curriculum, cross disciplinary skills, and pedagogy, as well as knowledge of learners and the community context.

Instructional Strategies

The teacher understands and uses a variety of instructional strategies to encourage learners to develop deep understanding of content areas and their connections, and to build skills to apply knowledge in meaningful ways.

Professional Learning and Ethical Practice

The teacher engages in ongoing professional learning and uses evidence to continually evaluate his/her practice, particularly the effects of his/her choices and actions on others (learners, families, other professionals, and the community), and adapts practice to meet the needs of each learner.

Leadership and Collaboration

The teacher seeks appropriate leadership roles and opportunities to take responsibility for student learning, to collaborate with learners, families, colleagues, other school professionals, and community members to ensure learner growth, and to advance the profession.

Appendix 2: Participant Self-Reported Demographic Information

Professional level upon entrance to program	Number	Percentage
Teacher	290	86.8
Administrator	16	4.8
Paraprofessional	6	1.8
Other	22	6.6

N=334. Two participants did not respond to this question.

Highest Degree Earned	Number	Percentage
Baccalaureate	110	32.7
Masters	209	62.2
Doctorate	8	2.4
Other	9	2.7

N=336.

Licensure Status	Number	Percentage
Certified	305	91.6
Not certified	12	3.6
Provisional	7	2.1
Emergency	9	2.7

N=333. Three participants did not respond to this question.

Years of Teaching Experience	Number	Percentage
0-2	51	15.4
3-5	76	23.0
6-10	62	18.7
11-15	68	20.5
16-20	38	11.5
21-30	31	9.4
31 or more	5	1.5

N=331. Five participants did not respond to this item.

Where do you teach	Number	Percentage
Public school district/charter	317	95.2
Private school	10	3.0
Not currently teaching	6	1.8

N=333. Three participants did not respond to this item.

Purpose of participating in ITQ	Number	Percentage
Professional development	318	97.0
Other	10	3.0

N=328. Eight participants did not respond to this item.

Race/Ethnicity	Number	Percentage
Black, non-Hispanic	41	12.3
White, non-Hispanic	264	79.0
Hispanic	15	4.5
Asian/Pacific Islander	5	1.5
Other ^a	9	2.7

N=334. Two participants did not provide responses to this item.

^a Six participants indicated that they are biracial or multi-racial.

Gender	Number	Percentage
Male	40	11.9
Female	296	88.1

N=336.

Did this activity take into account the needs of historically underrepresented and underserved students/groups?

Response	Number	Percentage
Yes	166	49.7
No	40	12.0
Not sure	128	38.3

N=334. Two participants did not respond to this item.

Note: two questions on the Demographic Form are not reported because so many respondents skipped or did not understand the question. These are #5 (Number of hours beyond Baccalaureate degree) and #10 (Certification area). Since most participants were elementary teachers, the options in #10 did not apply. We recommend that DDoE revise these survey items for more complete reporting.

Appendix 3: Sample Participant Survey

Title II Professional Development Teaching Non-Fiction Writing with Style – July 6 – 14, 2015

This survey is part of the evaluation of the University of Delaware's Title II professional development (PD) activities being conducted by the Delaware Education Research & Development Center. All participants in the PD institute are being asked to complete it. The survey is voluntary: You may skip all the questions or any question that you do not want to answer. While completing the survey does not have any direct benefit to you, it will be used by the UD faculty and staff to make improvements to future PD activities. It will take 5-10 minutes to complete.

What grade level will you be teaching in 2015-16? Elementary Middle High Other

Directions: Please circle the response that best matches your opinion. Mark only one response for each item.

The institute:	Strongly Agree	Agree	Disagree	Strongly Disagree
1. Was engaging and kept my attention.	SA	A	D	SD
2. Included a good mix of activities.	SA	A	D	SD
3. Used materials that are relevant to my teaching situation.	SA	A	D	SD
4. Had knowledgeable facilitators/instructors.	SA	A	D	SD
5. Gave me sufficient opportunities to ask questions.	SA	A	D	SD
6. Encouraged meaningful collaboration with colleagues.	SA	A	D	SD
7. Tried to cover too much in the time available.	SA	A	D	SD
8. Modeled effective instruction.	SA	A	D	SD
9. Took place in a space that met our needs.	SA	A	D	SD
10. Addressed a topic that was new to me.	SA	A	D	SD
11. Was worth my time.	SA	A	D	SD
After participating in the workshop:	Strongly Agree	Agree	Disagree	Strongly Disagree
12. I learned new approaches for teaching students informative writing.	SA	A	D	SD
13. I better understand how to align my lessons to the CCSS.	SA	A	D	SD
14. I have seen examples of activities that would work for my students.	SA	A	D	SD
15. I received helpful feedback on my lesson plan as I developed it	SA	A	D	SD
16. Lack of resources at my school will make the strategies we covered difficult to implement.	SA	A	D	SD
17. I feel confident in my ability to implement ideas and practices from the workshop in the 2015-16 school year.	SA	A	D	SD
18. My writing lesson plan needs more work before it is ready to be used with students.	SA	A	D	SD
19. I would like to participate in follow-up PD to support my writing instruction.	SA	A	D	SD

20. How do you plan to incorporate informative writing in your curriculum? Please be as specific as possible.

21. How do you think your students would respond to the strategies demonstrated in this institute?

22. What did you like best about the institute?

23. What would you change about the institute?

24. What challenges, if any, do you anticipate in using what you have learned?

25. Other comments

Thank you for your time!

Appendix 4: Coaching Participant Interview Protocol

1. Please describe *what you did* with the Title II coach.
 - a. If there were multiple coaching visits, probe on the different activities that happened in each one.
 - b. If necessary, follow up to identify what specific instructional techniques were targeted.
2. Please describe your style of interaction with the coach.
 - a. Comfort level with having the coach in the classroom
 - b. Ability to ask the coach questions
 - c. Relevance of the coach's suggestions to your teaching context.
3. How did participating in the coaching help you think differently about your teaching? Did you develop an understanding of specific skills or techniques?
4. Now think about application to the classroom. How, if at all, have you implemented the strategies that you learned through this PD?
5. What strategies are working well with your students, and why?
6. Are there any strategies that are not working with your students or that you have not been able to implement? Why?
7. Now think about your school context. What factors within your school are facilitating your implementation of these strategies? Who or what is helping?
8. What challenges have you faced as you attempted to implement your new instructional strategies into your classroom? What are the barriers?
9. How are the new strategies impacting student learning in your classroom, and how do you know?
10. What support do you still need in this area?

Appendix 5: Administrator Interview Protocol

1. Please tell us how this PD came to be offered in your school
 - a. Probe: relationship with UD and the Title II activities
 - b. Establish whether the administrator was directly present during the PD or not.
2. What were your goals for having your teachers participate in this professional development? What need did it address? What did you hope or expect it would do to influence teaching and learning in your school?
3. Did this professional development have the effect you had expected? If not, how was it different?
4. How have teachers in your school been implementing the skills/strategies they were taught in the professional development? (Probe for examples.)
5. How have students been responding to the skills/strategies from the professional development? Do you see any impacts on student learning or engagement, and how do you know?
6. How well do the strategies taught in this professional development fit with the way you already (teach reading/build classroom community etc.) in your school? Please explain.
7. How have you supported teachers as they implemented what they learned in PD to their classrooms? What support do they still need?
8. What has been the most positive result of this professional development? (probe for specifics on instructional practice or student learning)
9. What suggestions would you offer the organizers for improving this professional development in the future?
10. Additional comments?

Appendix 6: Module Development Final Checklist

READING/Writing PROJECT FINAL CHECKLIST 2015

In no particular order...

From the Pilot:

- Annotate student work from the pilot; at least one high, one middle, and one from a SPED student. (See PPT to help with annotations).

Final edits to modules:

- Make any necessary revisions to the module based on feedback from the pilot.
- Be sure all lesson plans reference all attachments, consistently labeled throughout the lesson plan (Appendix A, B, C etc.) and consistently labeled on the actual attachment
- Check the title! Is it "right"? Does it adequately describe the module (and hopefully grab the reader's interest?)
- Check planning charts, performance assessment, and lessons to ensure they are all aligned; standards, progressions, DOKs.
- Make sure formative assessments are aligned to the progressions in each lesson plan. Make sure teacher notes reflect help with using the evidence of the formative assessments.
- Check all standards/progressions – were they actually and explicitly *taught*? Were the standards/progressions listed the primary focus of the instruction?
- Check all prerequisites – make sure all standards-based concepts in the lesson that were *not* explicitly taught are listed as a prerequisite. [Note: if an entire standard, at mastery level, is NOT a prerequisite, consider underlining which part/target of the standard is the actual prerequisite. We do not want teachers to think students must master an entire standard as a prerequisite, if that is not the case].
- Check the "graphic organizers," "teaching strategies," and "sources" sections to make sure they match what's in the lesson plan. Check the vocabulary to preview – was it actually taught? Correctly labeled Tier 2 or 3/
- Make sure the format for the instructional chunks, including teacher notes, is consistent across all lessons. In fact, we *strongly* suggest that teams go through *all* the lessons together to assure consistency (for example, make sure the format and word choices match).
- Make sure that ALL differentiation is embedded and that there are plenty of opportunities for choices in formats, products, materials, etc.
- ALL appendices/attachments are clearly and consistently labeled. Make sure what's in the lesson plan matches the order of the appendices attached. ALL appendices/attachments must follow each lesson (exception: links to active websites).
- Double check that all sources are correctly noted and have correct citations. Make sure all links "work."
- The following FOOTER must be placed on each page (along with PAGE #s)

To introduce the module:

- Overview/plan: Write a paragraph of introduction for the unit. Something that will give teachers a very brief glimpse of the module's purpose, the topic/content, and the final product.
- Submit your finalized PowerPoint (a final form of what you did for Literacy Cadre)

Appendix 7: Summer Writing Institute Rubric

7/5/2015

Your Rubric - Print View



Rubric Made Using:
RubiStar (<http://rubistar.4teachers.org>)

Delaware Writing Project 2015 Lesson Plan Rubric

Teacher Name: **Ms. Howton**

Student Name: _____

CATEGORY	Outstanding	Good	Acceptable	Needs Work
Standards	All CCSS being explicitly taught are listed and specific skills within the standards are identified.	All CCSS being taught are listed and at least some specific skills within the standards are identified.	All CCSS being taught are listed.	Standards are not identified beyond the most basic identification (WS2) or are not identified at all.
21st Century Literacy	Students are both using and learning 21st Century skills which are clearly identified in the lesson plan.	Students are both using and learning 21st Century skills but they may not be clearly identified in the lesson plan.	Students are either using or learning 21st Century skills which may or may not be clearly identified.	21st Century skills have not been incorporated into the lesson plan.
Differentiation/Scaffolding	Lesson plan clearly identifies opportunities and teaching strategies to differentiate and scaffold the lesson. Teaching strategies allow for personalization to address individual student interests and needs.	Lesson plan identifies opportunities and teaching strategies to differentiate and scaffold the lesson. Teaching strategies are effective and research based.	Lesson plan identifies opportunities and teaching strategies to differentiate and scaffold the lesson.	Lesson plan does not identify opportunities for differentiation or scaffolding.
Exemplary Teaching of Writing Practices	Lesson plan clearly incorporates exemplary teaching of writing practices throughout the lesson plan.	Lesson plan incorporates one or more of the teaching of writing practices discussed during the class.	Lesson plan shows evidence of using the teaching of writing practices discussed during class.	The lesson plan fails to use any of the teaching of writing practices discussed during the class.
Mentor Text(s)	Lesson identifies specific mentor texts and incorporates them into exemplary teaching of writing	Lesson identifies specific mentor texts for use with specific activities.	Lesson identifies activities which use a mentor text and makes at least one suggestion for an	Lesson does not use mentor texts.

http://rubistar.4teachers.org/index.php?screen=PrintRubric&rubric_id=25403066

	practices.		appropriate text.	
Extended Thinking	Included in the lesson is an activity or assessment which requires students to use higher level thinking skills.	Included in the lesson is a suggestion for an extension activity or assessment which requires higher level thinking skills.	A supplemental list of additional activities or resources includes an activity or assessment using higher order thinking skills.	There are no mentions of activities or assessments requiring higher order thinking.
Final Writing Task	The final summative writing task reflects best practices for the teaching of writing and is a natural final product based on the lesson. The activity is a part of the writing process and students receive clear directions on the expectations for the product.	The final summative writing task reflects best practices for the teaching of writing and is clearly connected to the rest of the lesson. The activity is at least loosely connected to a part of the writing process.	The final summative writing task reflects best practices for the teaching of writing.	The final summative writing task does not reflect best practices for the teaching of writing.



Appendix B

ARTIFACT B: STUDY OF THE TEACHER LEADERSHIP INITIATIVE

Executive Summary

This qualitative study of the Teacher Leadership Initiative (TLI) by the Delaware Education Research & Development Center (DERDC) emerged from conversations between program managers and researchers. TLI is a program of Schools That Lead (STL), a non-profit organization in Delaware. DERDC is an independent, university-based center. In spring 2016, STL and DERDC leaders began to discuss the possibility of collaborating to study TLI. At that time, TLI was completing its second year. Program leaders sought to understand the early impacts of the program, in order to (1) make improvements for future cohorts and (2) potentially provide data to seek continued support and funding. Meanwhile, DERDC saw an opportunity to leverage staff interests and expertise and to inform the state educational system as it developed teacher leadership and professional learning initiatives. DERDC receives annual funding from the Delaware Department of Education (DDoE), which supports projects mutually agreed upon by DERDC and DDoE leaders. In 2016-17, DERDC proposed to use state funding for the current study, and DDoE approved.

TLI brings together cohorts of participating teachers for two years of professional learning including 14 full-day sessions, eight in the first year and six in the second. In addition, teachers are expected to conduct application activities in their schools. The purpose of the TLI is to: “increase student achievement through engaging teachers in pursuit of their goals for instruction, with efficiency to support

sustainability and scaling” (Schools That Lead, n.d.). The program is essentially structured as an inquiry project, where participating teachers pose questions about their students’ learning, then work with peers to collect data to investigate these questions and inform their practice.

The study addressed the following six questions:

1. Who gets involved in TLI, how, and why?
2. How do participants describe their experiences within TLI?
3. How do relationships between TLI teachers, other peers, and administrators develop during the TLI?
4. How does participating in TLI influence the instructional practice of participants? Others?
5. What impacts are evident on student learning?
6. What organizational conditions and supports facilitate or constrain the work of TLI participants?

Findings are discussed in depth and with many exemplars for each question, with #4 and 5 combined because issues of teaching and learning were so intertwined in participants’ narratives.

The primary method of data collection was interviews, along with some observations and review of project data and documentation. In total, 27 interviews were conducted in winter/spring 2017: 17 teachers, eight school administrators and two district administrators. The sample included individuals in the eight schools that had at least one teacher who completed cohort 1. These are located throughout the state and represent a variety of school types: elementary, middle and high; regular public, charter and magnet. The interview sample provides an in-depth cross section of the first two cohorts of TLI participants and their administrators.

A major theme in the data was that TLI participants had to “want to do it.” Cohort 1 in particular tended to be experienced teachers who already held some influence and actively pursued professional learning. There was consensus that the selection of participants for TLI was key. Teachers expressed goals for TLI related to strengthening both their instruction and their leadership; administrator goals mostly focused on the latter. We found strong satisfaction with the TLI professional learning days – content, facilitation, and collaborative environment. We also found that most participants brought TLI back to their school settings by completing the Student Learning Reflection Cycle (SLRC), although not all implemented it fully, and all faced time and scheduling barriers. Those who pursued the SLRC furthest provided specific examples of what they learned about student learning, and how they shifted their instruction in response. More broadly, we found that most TLI teachers described thinking about their classrooms in new ways. Their administrators also generally reported growth in their confidence and leadership.

Trusting relationships were integral to successful implementation of the TLI. Most teachers initially leveraged existing colleague relationships; some of these deepened through this joint work. In some but not all schools, TLI collaboration was starting to move beyond the initial group. We examined the spread of TLI in each school, finding significant variation in scope and pattern, and mixed feelings about the process of scaling. The program has a dynamic tension between fostering a small community of “likeminded people” and moving beyond that core to try to change schools. The story of TLI in each school is highly contextual and depends on factors including professional culture, administrative leadership and resources of time and funding. For culture and leadership, we found examples of both facilitators and

barriers in our data. Time presented challenges across the board, though some schools got creative with scheduling and/or through administrator support were able to prioritize TLI. Financial constraints have already been a challenge to implementation in some cases, and future issues are anticipated. Overall we found that building administrators played a crucial role in TLI and had to strike a delicate balance between being “hands off” and providing necessary support and alignment.

Introduction

This qualitative study of the Teacher Leadership Initiative (TLI) by the Delaware Education Research & Development Center (DERDC) emerged from conversations between program managers and researchers. TLI is a program of Schools That Lead (STL), a non-profit organization in Delaware. DERDC is an independent, university-based center. In spring 2016, STL and DERDC leaders began to discuss the possibility of collaborating to study TLI. At that time, TLI was completing its second year. Program leaders sought to understand the early impacts of the program, in order to (1) make improvements for future cohorts and (2) potentially provide data to seek continued support and funding. Meanwhile, DERDC saw an opportunity to leverage staff interests and expertise and to inform the state educational system as it developed teacher leadership and professional learning initiatives. DERDC receives annual funding from the Delaware Department of Education (DDoE), which supports projects mutually agreed upon by DERDC and DDoE leaders. In 2016-17, DERDC proposed to use state funding for the current study, and DDoE approved.

Program context

There is a growing interest in teacher leadership as a catalyst for educational improvement. In 2011, the Teacher Leadership Exploratory Commission released model standards for teacher leaders, whom they defined thus: “a teacher who assumes formally or informally one or more of a wide array of leadership roles to support school and student success...Teacher leaders model continual improvement, demonstrate lifelong learning, and use what they learn to help students achieve” (p. 37). Concurrent and intersecting with this focus on teacher leadership is growing

innovation in professional learning. New models have emerged for ongoing, job-embedded learning instead of traditional “sit and get” training. These trends are evident on both national and state levels. In 2012, Delaware adopted the Learning Forward Standards for Professional Learning. In 2017, DDoE demonstrated its continued focus on professional learning through this definition:

Professional learning is defined as a comprehensive, sustained, and intensive approach to improving teachers’ and principals’ effectiveness in raising student achievement. There is a shift from the concept of professional development to professional learning to connote the importance of continuous improvement (DDoE, 2017)

Meanwhile, teacher leadership has also been an ongoing state focus. In 2016, DDoE launched a Teacher Leader Pilot, which developed five new teacher leader roles in a small group of schools with the goal of “creating more career opportunities for Delaware educators while leveraging their talents to support students” (DDoE, 2016). There are several other initiatives for teacher leaders in the state, including the TLI. The program thus fits within the broader thrust towards enhanced professional learning and teacher leadership opportunities, but it uses a unique approach to accomplish these aims. TLI was launched in 2014, although STL and its leaders had been active in the state before that. The program is now funded by national and state organizations and foundations and by tuition paid by participating schools/districts. There was no tuition/cost for Cohort 1 teachers in the first year (2014-15). In the second year, the program “came at the cost of bringing a colleague to start what became Cohort 2,” according to STL staff.

Program design

TLI brings together cohorts of participating teachers for two years of professional learning including 14 full-day sessions, eight in the first year and six in the second. In addition, teachers are expected to conduct application activities in their schools. The purpose of the TLI is to: **“increase student achievement through engaging teachers in pursuit of their goals for instruction, with efficiency to support sustainability and scaling”** (Schools That Lead, n.d.). The program is essentially structured as an inquiry project, where participating teachers pose questions about their students’ learning, then work with peers to collect data to investigate these questions and inform their practice. This process is known as a Student Learning Reflection Cycle (SLRC) and will be detailed more below.

In 2016, STL developed a system of micro-credentials that, although optional for participants to pursue, maps out TLI activities. The 15 micro-credentials are organized into three categories: **Advancing powerful student learning in my classroom, Advancing powerful student learning in peers’ classrooms, and Scaling powerful student learning in my school.** As this language shows, the focus of the program begins with the individual participant and classroom, then “zooms out” over time. By year 2, participants are asked to begin scaling TLI activities and involving more colleagues. The scaling model is discussed in more detail below.

TLI has enrolled three cohorts: 1 (2014 – 2016), 2 (2015 – 2017) and 3 (2016 – ongoing, expected completion in 2018). The program continually develops as lessons are learned; according to STL, Cohort 1 experienced a curriculum that has evolved since 2014. Each cohort is comprised of teachers from different kinds of public schools across the state. Often, pairs of teachers from the same school attend together. In addition to teachers, school administrators are invited to attend TLI activities,

particularly a planning day held in winter of year 2, when participants develop the scaling plan/process for the school. In 2013-15, STL also offered a separate program, the Principal Leadership Initiative (PLI), which aimed to build administrators' skills in instructional leadership and feedback. The PLI is not part of our study, although some participants also attended it.

Study design

The purpose of this study was to qualitatively study the experiences and document the impacts of the TLI on the first two cohorts of participants, and to generate recommendations to inform future research and other initiatives related to teacher leadership and professional learning in Delaware. The study addressed the following six questions:

1. Who gets involved in TLI, how, and why?
2. How do participants describe their experiences within TLI?
3. How do relationships between TLI teachers, other peers, and administrators develop during the TLI?
4. How does participating in TLI influence the instructional practice of participants? Others?
5. What impacts are evident on student learning?
6. What organizational conditions and supports facilitate or constrain the work of TLI participants?

Methods

The primary method of data collection was interviews, along with some observations and review of project data and documentation. Data collection procedures were approved by the University of Delaware's (UD) Institutional Review Board.

Schools That Lead (STL) provided DERDC with contact information for members of Cohort 1 and Cohort 2. We determined that veterans would provide the most meaningful data about program impacts. Thus this study included participants and administrators from the eight schools that had at least one teacher who (a) was enrolled in TLI's first cohort and (b) completed two years of training. These schools represented five districts or charter schools in Delaware. The selection criteria mean that not all Cohort 1 and 2 schools and participants are included in this study.

According to STL records, teachers from nine additional schools attended some Cohort 1 activities but did not complete the program. TLI indicated various reasons for this attrition, including staff departures and funding decisions (i.e., school/district would not fund year 2). In addition, four schools joined the program in Cohort 2. These teachers and schools were not included in this study.

STL communicated the study and its purpose with Cohort 1 and 2 participants in winter 2017. DERDC then invited participants meeting our criteria in waves, first Cohort 1 and administrators, then Cohort 2. Table 2 shows the response rates. This sample provides a representative cross-section of TLI. It includes at least one Cohort 1 teacher and administrator from all eight schools meeting our criteria, and Cohort 2 teachers from six. The total number of interviews per school was two to four. In addition, we interviewed two district administrators who, although not listed in TLI records, offered useful perspectives.

Participating schools are located in all three counties in Delaware and include regular, charter, and magnet schools. Three are elementary (grades K-5 or 1-5), two are middle (6-8), one is middle/high (6-12) and two are high schools (9-12). Their level of participation in TLI at the time of the interview (winter/spring 2017) varied.

Five schools had teachers in Cohorts 1, 2 and 3. One school had participants in Cohorts 1 and 2 but not 3. One had participants in Cohorts 1 and 3 but due to staff departure, not Cohort 2. Finally, one school identified as no longer participating in TLI as of the 2016-17 school year; it had teachers in Cohort 1 and 2, but the latter only completed one year of training. Issues of attrition are discussed below.

Interview protocols are found in Appendix 1; there were separate versions for teachers and administrators. With one exception, interviews were held face-to-face.⁴ They ranged in length from 20 minutes to an hour but averaged about 40 minutes. Interviews were digitally recorded, transcribed, and coded qualitatively for themes related to the six study questions. Dedoose software was used for analysis (SocioCultural Research Consultants LLC, 2016).

In addition to the interviews, a DERDC researcher observed a planning day for Cohort 2 (January 2017). To build our understanding, we attended TLI activities before the official start of this study (Cohort 2 Year 1 Training, March 2016; All Schools Conference, May 2016). We held discussions with STL staff and reviewed program documentation including agendas, handouts, and the website, as well as participation and micro-credentialing data provided by STL. Finally, we asked participants to complete a form identifying colleagues with whom they worked in TLI; these data were analyzed to understand patterns of scale and participation.

Table B1 Study participation and response rate

Group	Number participants ^a	Total number	Response rate
Cohort 1 Teachers	10	11	90.9%
Cohort 2 Teachers	7	10	70.0%
Principals/Aps	8	8	100.0%
District contacts	2	N/A	N/A

Findings

1. Who gets involved in TLI, how, and why?

Most teachers in Cohort 1 described receiving communication about the program via email. Some initiated the idea of TLI, bringing it to their administrators. In other cases, administrators intervened, either by selecting a group of teachers to forward the STL email to, or by directly encouraging applicants. Some schools had already participated in the PLI, or other STL-led programs, and wanted to continue the relationship. The program emphasizes its voluntary nature, and indeed we found that all interviewees chose TLI but they often had their principals' voices and encouragement in their heads.

There was consensus that choosing the right participants for the program was key. Teachers and administrators alike described the voluntary nature of TLI as fundamental.

It's got to be grassroots. It has to be teachers that are interested in doing it, because if they're not, then it's not going to happen. They'll say that it's happening, but it's really not.

You can't make those kinds of things mandatory, so you're always dealing with people who care about kids and care about becoming a better professional who are going to take advantage of that.

As this second quotation demonstrates, TLI teachers were people who “cared” and were motivated. As one put it, “*I am a “yes” person*”; a couple talked about their own “*growth mindset*.” Thus some participants gravitated towards TLI simply because it was a professional learning opportunity, and they wanted to grow. The majority were experienced teachers, with an average of more than ten years of experience in their current schools (i.e., the ones joining TLI) and, in some cases, other prior assignments. Almost all reported currently or previously holding leadership

positions such as department chair, leadership team member, mentor/instructional coach, assistant athletic director, or union representative. Over their careers, most had been active in other professional learning initiatives, including some with state or national scope (e.g. National Board for Professional Learning Standards, Delaware Math Coalition). Despite this wealth of prior experience, many teachers emphasized how different TLI felt to them:

There's not really a lot of training around what good teaching looks like, and what good learning looks like. And so that was kind of what my interest was around, like 'Why can't we do some things on our own to facilitate discussions around what good teaching and learning looks like within the building?'

Participants saw that TLI was focused on student learning and that it was teacher driven (“do some things on our own”). These ideas will be explored throughout this report.

What piqued teachers’ interests about TLI specifically? Most Cohort 2 teachers heard about the program through Cohort 1, and some had already participated in program activities such as the SLRC. On the other hand, Cohort 1 teachers joined a brand-new program about which few details were initially available. Some wanted to re-invigorate or to take their teaching to the next level in ways that current PD or evaluation systems were not accomplishing.

So it just intrigued me what could I do differently after teaching for 20 years, what can you do differently?

Others articulated goals related to teacher leadership or future career goals:

I needed to learn about how to be a more powerful leader.

“Maybe I’ll learn something about how others perform their craft in the classroom so it will help me become a better administrator.” That was like sort of my motivation.

Cohort 1's goals tended to be fairly general because of the newness of the program.

Since many administrators played an important role in recruiting participants, their goals are also examined. For Cohort 1, they primarily also wanted to strengthen teacher leadership. These comments provide an early look at what administrators meant by that, and what they expected from the program upon entry:

[The goal was] taking teachers, good teachers, and giving them skills to be teacher leaders within a building and to feel comfortable with their practice to share it with other teachers.

I felt like we had already been doing [teacher leadership] and it would benefit us to send a teacher to that just to kind of bring back some more of what that's looking like other places and share what we're doing here.

Administrators also viewed the choice of participants as important. In general, Cohort 1 teachers were seen as influential players in their schools, and this helped propel and scale the program.

By pulling in the right people, and getting their buy-in, and by making it people who are very trusted by the staff, that was our first big win.

Advice for a fellow administrator: Be very careful on who they select because that person can be very powerful with their message that they bring back.

So when they saw those two teachers especially really liking this process and showing what it could do, it allowed the other staff to see that it wasn't going to be bad.

Finally, some administrators were enthusiastic at the prospect of the TLI because they had had positive experiences in the PLI or other prior STL programs.

2. How do participants describe their experiences within TLI?

This section examines two aspects of the TLI experience. It briefly describes participant response to the professional learning days and then focuses on

implementation, i.e. how participants used and applied work from TLI back in their schools. We look especially at the Student Learning Reflection Cycle (SLRC). Subsequent aspects of the work (e.g., instructional shifts, scaling) are considered in later sections.

TLI professional learning days

Although the interview protocol did not ask specifically about satisfaction with TLI days, many participants brought it up. Several stated this was the best professional learning experience they had ever attended. One teacher described it as more powerful than anything else because it was collaborative and sustained over time:

You have the ongoing support and the continual talk and the continual reinforcement, and it's not a one and done PD.

Others praised the STL facilitators for “walking the talk,” modeling how to get teachers to reflect in detailed, rigorous ways about practice:

Sofi and Dana are so engaging and enthusiastic and passionate about what they do that it's very contagious, and it's easy just to keep going and participating.

They were doing to us what we should be doing to our students. It was like just questioning us and making us work and like not accepting like a blank stare at something, right?

Some contrasted TLI with more passive prior PD experiences. One talked about bringing a stack of papers to the first session in case it was boring. S/he never got to that grading.

Most participants also found the TLI community motivational. We heard words like “excited” and “giddy” to describe the feeling of being in this company. One commented,

It was just kind of refreshing to be on a cutting edge of something with people in a positive environment

In describing their TLI peers, many participants used the phrase “likeminded people.” This is one of several examples in our data where vocabulary recurred, suggesting that participants internalized some key TLI messages. Later in this report, we discuss the implications and the potential limitations of this focus on being “likeminded.” Participants clearly valued the peer environment in TLI. They appreciated the opportunity for statewide and vertical networking. Several identified this as a benefit they could not get from school or district-based PD.

I know that that piece of the puzzle's really powerful when you are not sitting and listening to the math supervisor, who you've seen for the last eight years, tell you the same thing and you're still not doing it. You're going somewhere else, "What? You guys say do this too? Maybe I will try it."

You feel like you're on an island anyway, so having conversations with other schools who were working to improve teacher effectiveness was great. So just that building of relationships was huge for [teachers], and that was probably the biggest thing that they brought back.

Most participants got a boost from spending the day at TLI. The greater challenge came when they returned to their schools to do their homework and apply what they had learned. This took greater persistence and patience.

When I'm in it, when I'm there, it's great. I love it. It's great ideas. It's inspirational. When I'm not in it, it's like all my other stuff. So I have to do a better job of making those things a priority I would say. I think that's challenging

You go to a program like this. It's organized. It's got great ideas. [Participating teachers] said as much every time they'd come back. "How was your day?" "Oh, it was awesome." But then you do come back to your room, and teachers are busy, and you just can't implement everything you learn, and you certainly can't spread that word as quickly as you probably would like when you get in your car at 4:00 in the afternoon.

Finally, we note that not all comments about the TLI professional learning days were positive. Three teachers and one administrator raised at least some concerns. These related to pace or scope (TLI moved too slowly, covered things they already knew), application (TLI was too abstract) or focus (TLI should have emphasized structural or leadership issues more).

I just felt like I wasn't challenged enough like I wanted to be. I felt like it was more like work on my shoulders than people challenging me to become an expert in my craft...

I just felt like it was almost an oversaturation of the same content... in my mind, it was almost like the same lesson over and over.

Sometimes, initially, the first year, I never knew where the meeting was going.

A small number of participants may have felt themselves to be “ahead of the curve” in teacher leadership, or had a different understanding of what that role entailed, contributing to these comments. But they are a minority view. The last idea, that plans were sometimes not clear, connects to broader issues surrounding communicating TLI. These recur throughout the findings.

Student Learning Reflection Cycle

TLI also included numerous activities back in the schools. In this study, we focused on the Student Learning Reflection Cycle (SLRC), which was the central, but not the only, application activity of TLI. Other activities described include conducting student surveys and shadowing a student for a day According to Schools that Lead (n.d.):

In a Student Learning Reflection Cycle, peers: (1) use a protocol to help each other identify key questions about their students' learning (2) design ways to collect data about those questions (3) reflect on what they learn and determine appropriate next steps in their classrooms.

Of the 17 teachers interviewed, ten (59%) spoke about the SLRC in ways that clearly indicated they had conducted at least one cycle in their own classroom since beginning TLI, a span of between 1.5 and 2.5 years. They could tell the story in detail, e.g., identify the student learning question, the data collection process, and the major takeaways. Three (18%) described experiences in more general terms and four (23%) could not provide examples. Of these, one was planning to have a peer collect data in his/her room in the future; two cited scheduling and other issues that interfered; and one did not explain. Notably three of these four teachers had visited/collected data for a peer, although there was not yet reciprocity. As TLI implementation progressed, early participants more often found themselves in the role of visitor as they showed others the process. One went as far as to state that TLI is “definitely more of a selfless program because you are helping others more than you are looking at your own classroom.” Everyone agreed that the SLRC was harder than it looked. TLI scaffolded the process and provided practice opportunities within their sessions. For instance, the January 2017 planning day included a detailed discussion of the SLRC, a video of two Cohort 1 teachers in conversation, and a simulated observation with classroom footage. Whether as a host or visitor, interviewees found that the SLRC got easier over time. They improved at posing incisive questions, honing data collection instruments, and focusing their attention during the observation. As host teachers, they obtained data related to student experiences and learning, which provided them with opportunities for reflection (see questions 4 and 5, below). At its best, the SLRC allowed teachers to pause the action and think more deeply: So many of teachers’ decisions are spur of the moment decisions. I actually got to sit and reflect and think, “What do I want do with this now? Where do I want go next?”

Participants also described benefits of the visitor role, including seeing their students in a different context, getting instructional ideas, and developing their teacher leadership skills (e.g., talking with others about instruction).

Participants had clearly heard and internalized the message that the SLRC was non- evaluative, even if their non-TLI colleagues had more difficulty understanding that:

The last time [observation partner] was in here, s/he said, “Great job doing this.” And I went back to her. I’m like – I just put my hand over that part of the notes, and I said, “I appreciate it, but you just need to write down everything.” And s/he knows I’ve been through the [TLI] training. And s/he’s okay with that.

Holding back from judgment, whether positive or negative, was sometimes difficult but was essential to the integrity of the SLRC process. It could also be a relief:

It’s rare for a teacher, I think, to just totally release that feeling of responsibility or that feeling of control. I’m purely here to observe. So, I really enjoy that aspect of it.

Almost every time we heard about the benefits of the SLRC, we also heard about the challenges of scheduling and fitting it in. As one teacher commented on his/her difficulty executing the SLRC, *“even the fact that we find it extremely valuable, that fact doesn’t supersede time and space.”* Time barriers are discussed further in question 6, below.

3. How do relationships between TLI teachers, other peers, and their administrators evolve during the TLI?

Relationships within TLI: pairs and groups

Peer collaboration was integral to the TLI. Most teachers chose those with whom they were already close – “*comfort colleagues*,” as one put it – to involve in this new process. They also looked for partners who would be open to TLI.

Obviously, when you start this process you're going to go to your BFFs.

It takes one thing off the plate – that learning to work together piece, because it's already there.

They're people that I am close with, and so we are comfortable with each other and I think ... we will be honest and share our ideas and feedback and those sort of things without feeling like, okay, I don't feel comfortable with this person or what are they going to do with this data?

The initial collaboration helped the SLRC take root. Occasionally, it was not securely established, and implementation suffered. This seems to have been either because the participant did not choose his/her partner or because s/he worked with many people at the same time.

Asked whether TLI had nurtured their peer relationships, some participants disagreed because they had already been close: “*I would say no, just because of who I chose.*” Others described how existing relationships grew stronger, with more dialogue:

I think we're much closer because there's nothing like seeing you at your best and worst, and not worrying about it because it really wasn't about you. And you could have thought: “Man, that was really awful.” And they're saying: ‘Well, I have the data that says it wasn't because I wasn't paying attention to you. And did you realize that so and so said this, and so and so said that?’

Yeah, I think we're a lot closer. We go to each other all the time now. I think we have a different culture. We've always had a really good open door policy in [grade level], we go to each other and talk, but I think s/he's become much more of a thought partner for me in all aspects, not just like I have this learning question. It's just, "This isn't going well, what do you think? Let me talk this out to you." You create that kind of relationship.

In some schools, relationships also formed more broadly among TLI participants. Four schools held at least periodic group meetings; in two, these were held monthly after school or during morning homeroom. Participants elsewhere hoped to launch group meetings in the future or wished they could but identified scheduling barriers. Group meetings supported implementation by keeping teachers accountable to each other and to the SLRC process:

I think they important to keep – well, for one to keep the excitement going, and to keep, really to keep the program the way it was intended.

We're constantly talking about trust, and I feel that that trust is there between the [TLI participants]. But, that trust is not there yet with everybody in our building. So, it's not like we can go marching down the hallway, knocking on people's door saying, "Hey, I'm going to do this [SLRC] with you." There has to be that buy in.

Relationships outside TLI

As this quotation demonstrates, TLI teachers also interacted with other professionals in each school. Descriptions of these cultures varied widely, everything from “*highly collaborative*” to “*independent*” to “*probably unhealthy*.” The TLI program model contained a paradox: the work took root in relationships between pairs or small groups of “*likeminded people*” but aimed to change schools, and to do so it had to spread more widely. This participant articulated the risk of TLI being seen as a “*clique*”:

Sometimes I feel like....we're like a small secret society, and we're not. But to get other people to come, they're like "No, no, no. I'm not really – I'm not into that."

In some schools, TLI appeared to be a more open group than in others. Most described attempts to share out information about the program. Building administrators influenced official opportunities (e.g., time on meeting agendas) but informal conversations also occurred in some hallways and lunchrooms, within grade levels and departments. A cycle emerged: in schools where more teachers participated in TLI, there was more communication about it.

It's kind of like, I do TLI. And people are like, what is that? I heard so and so, and so and so do that. There's a little bit of a buzz.

However, in other schools we heard that TLI was still operating under the radar, that staff members only knew what was happening if they asked. When teachers outside the cohort learned about TLI, they had a variety of responses, from curiosity to pushback. We heard that, perhaps most often, teachers were so busy that TLI did not have much of an impact:

I don't think it's a negative reaction, I think that it's just lack of engagement. I think people see it as "one more thing."

Colleague responses reflect the broader professional culture, which is discussed in question 6.

Scaling and snowflakes

Beyond just discussing TLI, the program expects that teachers will eventually “scale” the program in their schools by getting other teachers to participate. The scale model is illustrated by a snowflake (Figure B1).

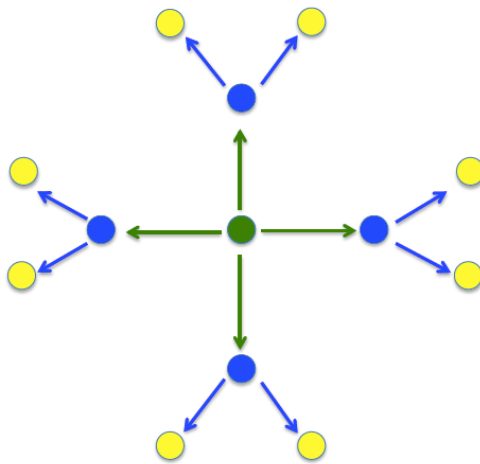


Figure B1 TLI snowflake scaling model

In this model (Stevens, Wallace, Frankowski & Mash, 2016), by the end of their second year, each cohort participant (“green dot”) would work with four other teachers (“blue dots”) who eventually would work with two others (“yellow dots”) to create a total group of 13. “Working with” each other primarily consisted of engaging in the SLRC together. Although blue and/or yellow dots might also choose to join a subsequent TLI cohort, this was not required.

Just as every snowflake has a unique shape, every school approached scaling differently and each was at a different stage of the process. At the time of the study (2.5 years after the program launch), one school in the sample of eight appeared to be approaching a full ‘snowflake’ with green, blue and yellow dots. While some growth was organic and emergent, driven by teacher interests, schools also had the opportunity to develop a scale plan. For instance, the Cohort 2 planning day was held in January 2017 and involved teams of administrators, teachers, and invited guests. STL staff shared materials about change management and circulated. to coach

teams and provide support. They cautioned attendees to remember the inter- and intra-personal dynamics of scaling and not get carried away with the technical plan. We observed strong engagement during this session but noted that not all teams were present. Two schools could not attend because of scheduling conflicts or personal circumstance; they planned to plan at their sites. No information is available about whether or how this occurred. We also know from Cohort 1 interviews that sometimes scaling plans materialized and other times they were adapted or abandoned.

Prior to the interview, teachers were asked to complete a form identifying others with whom they worked in TLI. On average, they listed just over four people (mean = 4.2). This slightly exceeded the program's expectation, but the extent of involvement varied, and it is unlikely that every individual listed operated as a "blue dot" (for instance, one participant wrote, "attended one event" next to a person's name). Also, responses ranged from one to six. Looking deeper into the data, we see more variation. In two schools, most or all of the collaborators listed were members of TLI Cohorts 1, 2 or 3. In other words, TLI was a fairly closed community. In two schools, the opposite was true: teachers *only* listed collaborators who were not TLI "green dots." In these cases, TLI teachers worked in parallel, each picking up his/her own group of collaborators. The rest of the schools used a fusion approach. While some secondary teachers focused within their content area, more often scaling occurred across disciplines or grade levels.

Another way to understand scale is to look at the number of teachers involved in TLI as a proportion of the staff size. "Two participants" means something different in a small elementary school or a large high school. At most, TLI appears to have touched about a quarter of teachers in a school. At least, it was isolated to a handful of

individuals on a staff of about 50. In one case, teachers and administrators wondered if the school was approaching saturation. Teachers in Cohort 1 and 2 had found their “blue dots” and some also had “yellow dots”, and Cohort 3 teachers were starting to think about scale. Although pleased with this momentum, participants raised questions about capacity and fidelity (e.g., what kind of monitoring is needed?). In this school, as well as in two others, demand for TLI exceeded available spaces and an internal application process had been developed. In two schools, however, questions about penetration or reach almost did not make sense, as the program was still trying to gain traction.

As TLI scaled, some saw participation diversifying. As might be expected, Cohort 1 tended to be self-starters and early adopters; as discussed above they were established within their buildings, and many already worked closely together. Later, there is some evidence that a more varied group of teachers was starting to buy in to and choose TLI. One administrator said:

The people that are starting to participate in TLI aren't necessarily the worker bees. It's people who are good teachers and have expressed an interest in getting involved in TLI, but they're not that core group that's always planning the PDs and giving the PDs and, you know, in a leadership role necessarily.

We heard various examples of TLI moving outside the “usual suspects” in some (though not all) of the schools. For instance, we learned of teacher getting involved who were: new to the profession or to their buildings; on the verge of retirement; people who formerly “*kind of did their own thing*”; or non-core/Related Arts teachers. In at least three schools, participants used the SLRC as a form of embedded teacher education for their student teachers, interns, or mentees. Still, we

also heard examples of groups of teachers who “would never” be interested in something like TLI. Another core belief TLI was, “you will never get everyone.”

One school stood as an outlier on this issue. This school is no longer officially participating in TLI; a collaborative decision was made to cease attendance halfway through Cohort 2. Essentially, the school believed that it had built internal capacity to manage a peer-to-peer observation process (an adaptation of the SLRC) itself. According to interviews, this school began to scale the program on a voluntary basis, reaching about half the teachers in the building in this way. Beginning in the 2016-17 school year, all teachers were matched up (with pairings made strategically by administrators) and expected to participate in a process they called “peer-to-peer walkthroughs.” Perspectives on the effectiveness of this approach, and the extent to which it reflected TLI practices and intentions, varied across the interviews from this school.

The scaling process surfaced challenges. Participants often felt that it was hard to explain TLI or the SLRC to their peers. One wished for better materials to illustrate and communicate the program. Further, many stated that “blue dots” had more difficulty executing the SLRC than “green dots.” They saw foundational understanding established through TLI professional learning as important, and it was difficult to catch people up in the limited school time available:

It took us two years to kind of feel comfortable actually listening to and getting an observation tool together. So for me to think I’m going to do that in two meetings or three is, I don’t know. It doesn’t make sense.

Discussions about the complexities of scaling were observed at the All Schools Conference (April 2016) as well as the Cohort 2 planning day (January 2017).

4. How does participating in TLI influence the instructional practice of participants? Others? and 5. What impacts are evident on student learning?

We consider these questions together since instruction and student learning were so intertwined in participants' narratives. As described in the introduction, TLI aimed to "advance powerful student learning" within participants' own classrooms, their peers' classrooms, and in the school as a whole. In this section, we first look at changes for participants, separating them into specific (i.e., related to the SLRC) and more general shifts. Then, we consider influences on peers and in the school as a whole. As we will show, these broader impacts are still emerging.

Specific changes in own classroom

Typically, teachers developed their student learning questions (SLQs) around an area they identified needed improvement. They also used STL tools (e.g., the Student Learning Rubric) to shape their questions. Often, questions pertained to times when the teacher was occupied with a subset of the class and an additional set of eyes would be particularly useful. For example, participants framed questions about students' depth of questioning or use of evidence in literature circles, academic discourse while studying math or vocabulary, or persistence or collaboration in small groups. Data collection often became quite specific. Some "zoomed in" to one small group (even, in one case, an individual), usually those who were struggling. Therefore, the data provided insights into how *particular* students were experiencing and learning (or not) particular content, at particular times of day.

In the SLRC, the visiting teacher leaves data collected with the host, who then decides what, if anything, to do about it. As they recounted their experiences, many teachers described making what one called "*little tweaks*." These were tightly related to the focus of the inquiry and included: re-group students, scaffold concepts, change

procedures (e.g., homework review, managing a discussion) to put more responsibility on students, use new formative assessments (e.g. response boards), refine questioning (e.g., higher depth of knowledge), define group or partner roles. Some teachers described collecting data repeatedly to track instructional/learning progress. (This was especially the case for teachers whose schools also participated in the Student Agency Improvement Community (SAIC). SAIC incorporates improvement science and a faster problem-solving cycle, which teachers can conduct independently. As participants generally saw TLI and SAIC as strongly aligned, sometimes they conflated activities for the two programs). In several situations, we heard that these instructional “tweaks” led to changes for students. For instance, one teacher gave students more choice in how to study a concept and saw their assessment scores rise. The long quotes below capture growth in student independence, participation and persistence.

*When I got the information, I basically could see that I was giving the prompting. And so I had to come up with a goal to set for myself, what can I do to decrease my prompting, but have them still doing it? So I let them rely on their peers a little bit more. So when I asked a question it was more to the peers, and they had to do it together. That’s kind of the conclusion that I came to, you know, from working together I can decrease their teacher dependence and increase their independence.
(Elementary teacher)*

What we found was that we were calling on the same 8-12 students in each class while the remaining 25 or so remained quiet or participated on occasion. From there we decided to load all of our students onto Class Dojo and click the random button. If their name came up they had to participate and we would give them guidance if needed. We did this for about two weeks where students stopped raising their hands and participated when their name was selected. We found that when we stopped doing this more students continued to participate because the norm had become that they had to be ready at any time. We still go back to Class Dojo about once a week to catch the students off guard

and make sure they are participating; however, we are now up to 20+ students actively participating on most days. (Middle school teacher)

What helped me that I could point to here was increasing the level of depth of knowledge. I mean, the first questions, they were too easy, so [students are] not going to persevere....The other question's probably more relatable, and more involved. So, they really had to work to figure out. And they found that more worth doing. So, that's what I found from – that was my big ah-ha last year. Just something that I talk about all the time with my department. Because I'm able to share this with my department. You know, that really, when they say that increasing the DOK, will increase [performance in] classes like this, because my teachers know some of these kids, too. And to say, it really did, here's proof. And, I tell people that all the time. (High school teacher)

Broader changes in own classroom

Participants emphasized that the SLRC focuses on students and that the data collector hardly notices the adult in the room. More broadly, this shifting focus from teaching to learning permeated our data. Reflecting on how their classroom had changed over the past few years, participants described a more student-centered learning environment, in which students actively engage with the material and with each other. This contrasts with a traditional “stand and deliver”, teacher-directed classroom. This theme was evident in almost all interviews across both cohorts. Here, one teacher from Cohort 1 describes how both s/he and her students now think differently about the class:

Before, I was very focused on teaching. And I've always felt – and the kids have always told me – that I'm a good teacher. I'm very good at breaking things down and showing them how things work. And to them – to students -- that makes me a teacher. But I wasn't getting them to think. You know? And [TLI] really made me open my eyes to that. Like, it's not...It's learning for September to June but carrying it with them somewhere else or making connections – that's not happening.

We often heard the idea that *“my voice has become less obvious in the classroom.”* One teacher pointed to a podium s/he used to use that was now *“in the back room collecting dust.”* Several teachers stated the difficulty of giving more classroom responsibility to students.

If I’ve got anything from TLI, it was that it shifted my focus as a teacher to the students. And I think that’s still challenging, too, because, as a teacher, you have your curriculum, and you have your lesson, and you want to cover material. And you have time constraints, and you have behavior issues. And you have a lot of different things going on in the classroom... And the challenge really is, are my students learning what I want them to learn? And how do they prove that to me?

Sometimes contextual situations (e.g., student needs or behaviors, teaching loads) interfered with instructional innovation. One participant acknowledged, *“I have to say that I’m not fully evolved into the teacher I want to be or [what] a TLI teacher probably should be. But there are times when I’m able to really use things that I’ve learned.”* Despite the challenges, this teacher persists, and communicates a sense of instructional efficacy. We heard that from others, too:

I find myself thinking about how I can change things more within my room as compared to just hoping that it will work out. I am more focused on how my students are learning and interacting with material than what their grades are. Overall, I find myself questioning what I am doing, what the students are doing, is this beneficial, and how it is impacting their grades.

In the situations where implementation appeared deepest, we were struck by parallels between what seemed to be happening for teachers and for their students: growth in confidence, willingness to try new things and make mistakes.

Because I often will say, “Well,” you know, “I’m trying this. Let’s try this new thing. We’ll see how it works.” You know? “Oops – that didn’t work. Okay, so let’s try it this.” You know? So that [students] see me take risks, they’re willing to take risks.

One of the biggest benefits for me, because I consider myself to be somebody who is a perfectionist to the point that I'm really hard on myself, I'm less hard on myself. So, I've given myself this grace period of, "Oh, guess what? You're not a perfect teacher and your students aren't perfect learners. But, you can reflect on why that is, and you can implement these small changes.

Data about instructional change were more pronounced in teacher interviews. Administrators tended to describe holistic growth rather than specifics, but we did hear some confirmation of classroom shifts. One described how a Cohort 1 participant “*led a very teacher- driven classroom*” before TLI but began to talk less and push students to deeper discourse. This principal described observing this teacher give “*the best lesson I've seen on [topic] to this day*” and credited that shift to TLI.

However, administrators also acknowledged that TLI participants were already strong teachers. While eight of ten identified some level of instructional improvement, there were some qualifications:

People that generally want to grow this way are pretty darn good to start with. I can't answer [the question about instructional change].

This caveat connects to a broader tension within TLI. Many of its first participants were also “high flyers” who already taught effectively and were motivated to improve. Greater challenges came as these individuals began to try to influence instructional practices of their peers and, eventually, their whole schools.

Influence on peers and school

TLI also challenged teachers to think more broadly about learning in their buildings, requiring them to build confidence as leaders and learn to influence their peers. Many were initially attracted to the “leadership” part of TLI even if they were not sure what it entailed. Over time some found that TLI gave them new access to relationships and instruction throughout the building. Only a couple of teachers gained

an official new role, but others found increased visibility as they presented the program in staff meetings, state or even national conferences. Just as they were mostly seen as strong teachers prior to TLI, cohort members were already viewed as influential. Still, the quotes below capture continued leadership development in the eyes of their administrators:

Our goal of pulling in [name] to help transform that team was 100 percent achieved. [S/he] definitely functions very differently now than [s/he] ever had. And is very much a participating member of the leadership team and really does work to drive change.

I've seen [participants] grow in their willingness and comfort level to reach out to teachers and simple things like that, like willingness to send an email to an entire group of people, saying, "Here's what I'm looking to try," to collecting feedback, working with other adults. I mean, that's a major shift to work [with adults] – to be good in front of kids is different.

I think they start to build a confidence in themselves because now they are learning how to improve themselves. So I think then they're like; okay, well I could probably help some other people.

How could TLI participants “*help some other people*” instructionally? One way would be to involve them in the SLRC as a host and/or visiting teacher, which in turn would give them the opportunity to learn and grow from that experience.

According to STL staff, as of May 2017 45 micro-credentials had been earned but only one of these was in the third category (“Scaling powerful student learning in my school. This suggests that broader instructional impacts were still emerging at the time of the study, 2.5 years after the program launch. When we talked with participants about instructional impacts, we also sought to understand if they saw changes in their *peers*’ teaching. We heard occasional examples about peers who were in TLI, whether in the same or a later cohort. Only a couple of administrators

discussed changes for teachers further out on the snowflake (i.e., “blue” or “yellow” dots, or people who participated in some aspects of the SLRC). The comments also tended to describe more general shifts rather than specific instructional changes. For instance, one noticed that his/her staff was initiating requests to TLI participants to visit their classes:

That’s a huge turning point. And I think it also shows that there’s a level of trust that’s beginning to develop between the TLI cohort and the rest of the staff.

This comment hints at professional culture, a topic we examine more fully in the next section.

6. What organizational conditions and supports facilitate or constrain the work of TLI participants?

As we have explored, most participants responded positively to TLI but had different degrees of challenge bringing practices back in their schools. In this final section, we look at some of the organizational factors that helped or got in the way. We group them into four categories: professional culture (especially as it relates to peer observations), administrative leadership, time resources, and financial resources. These four factors are highly inter-related, and they also connect to many of the findings already presented.

Professional culture

Teachers and administrators were asked to describe how adults work together. Responses varied. Some schools were described as having stable staffs who knew each other well, generally trusted and supported each other, and had collaborative routines (e.g., PLCs). This was more common at the elementary level. Even in these cases,

collaboration could vary by grade level or be vulnerable to the press of other obligations.

Overall, we get along really well. Everybody's – it's just the culture of education has changed so much that we're all so busy during the day that I feel like we don't interact as much as we used to ... We're all, everybody is working constantly.

Others reported mistrust or lack of communication between grade levels or between teachers and administrators.

It's pretty toxic, I would say, in this building. There are some groups that are very negative just towards – in general – being here in school.

Some of these tensions seemed to be historical; others related to decision-making or personal style. At the secondary level, departmental “silos” persist.

For the most part, they work really well in departments. But they are still high school teachers and they don't want people in their space.

In at least two schools, the professional culture was in transition, with significant turnover. In general, new teachers were seen as open to collaboration and learning from their peers, but this churn added to the dynamic context in which TLI operated.

Another aspect of professional culture is attitudes about sharing one's practice (deprivatization). Most schools had explored some kind of peer observation or walkthrough prior to getting involved in TLI, and some were still using it alongside the SLRC.⁹ Some administrators felt that provided a foundation for the SLRC and/or that the processes were similar or synonymous. Among teachers, a more prevalent view was that SLRC was different, and better, than other peer observations. Some brought negative prior experiences of peer observations rife with social or bureaucratic expectations. More than one person called them a “dog and pony show.”

It was all fluff....it wasn't authentic, it didn't help me at all become a better teacher. It just was, "oh, okay, this is great."

The SLRC was seen as more structured and meaningful because it was based off a teacher's own inquiry about his/her practice, and it took place in the context of a sustained relationship. Still, some teachers found they encountered some "baggage" from their peers.

I think because of the culture of teaching, to use the word "observation" and go into someone else's classroom, that's not something that's well accepted. It's hard to use the right words to say, "Let's just go into each other's classrooms and collect data."

This comment also reflects a theme evident throughout our data, that TLI was hard to explain to external audiences and made sense only in the application.

TLI took root more easily where there was already trust and collaboration among teachers, but it could also help schools improve staff culture. In three schools, we heard how this was a goal for the program:

And I'm hoping that [TLI] starts to build relationships among the staff – right now [they] are very cliquy...So I think it would just make it a happier place to be – more exciting. I knew it might take a little while but I felt like it would help with the instruction and eventually get to the heart of really teachers helping teachers get better.

In most but not all cases, TLI appeared to be having a positive cultural ripple effect, even if substantial challenges remained.

*I feel like the culture of our school in general is starting to change. I feel like it's much more open door. I can come in and talk to you and we're all connected. We sit down and it's nice to have – not just what's happening in my classroom but in our school. I feel like there's a shift. It's not happening everywhere, but like Sofi and Dana say, "You can't get everybody," but I think that if you can keep pulling in people and keep getting more people involved, it is fair to think that it's started.
(Teacher)*

Administrative leadership

Intersecting with teacher leadership, another important organizational factor was administrator (principal/AP) leadership. For TLI to flourish, administrative leaders had to make room for it. All eight interviewed described themselves as supportive of teacher leadership and convinced that it would improve their schools, yet they spoke about it in different ways. Here are two examples:

If you do [teacher leadership] right, you can free yourself up to do some other things, and you're not constantly trying to initiate programs. And, secondly, people are going to go the extra mile for you if they know that they've got a stake in the school and the school's decision-making and the school's future.

The only way I could continue to truly sustain work in a school would be to have it be teacher driven. So I looked at TLI as a way to create that teacher-driven culture. If I could send somebody to [TLI], then I could sort of create this group of likeminded people that would be a strong influence in the school.

How administrators positioned themselves relative to TLI varied, as did their own involvement in the program. In five schools, administrators were active in the PLI and/or other STL-run programs; in one school, there was sporadic involvement; and in two schools, little involvement. We also note that one teacher from Cohort 1 is now an administrator in a different school within the same district. Administrators' background knowledge about the program was an important facilitator:

I think that connection [between PLI and TLI] was really good that first year for me because I understood what the heck was going on.

In general, the administrators who had a deeper understanding of TLI through their own participation in related activities spoke in more detail about the ways they supported the process. Some also talked about their own belief in the value of the program:

I felt it and tasted the Kool-Aid, and it was like: oh, wow, this is really getting people to think differently about what we do every day that we think we know so well, and challenge it. Whoa, like, that's uncomfortable, but here's the growth.

On the other hand, some administrators saw their main role providing teachers with time and funding to attend:

With any professional growth opportunity, I think my role is to make it happen.

From this perspective, the fact that some of their staff was in TLI did not necessarily affect a principal's own work back in the school.

Many administrators felt they could help TLI teachers have a platform for communication and influence within the school:

I think my role is to be in the background and encourage them and support them, but also to help them spread the word so that we can build our capacity here in the building.

This looked like giving TLI time/space on faculty meeting agendas as well as providing coverage, supporting substitutes, supporting travel to conferences, etc. Some administrators also tried to communicate support for the program, while not pressuring anyone to participate:

My job was to really share out, share out, share out. You don't have to do this but this is what we're doing. You don't have to do this but this is what we're doing.

Like the teachers, TLI administrators had clearly internalized some program messages, including the idea that they needed to be "hands off." Discussing participants' choice of student learning questions, one administrator said,

Even if they're focusing on how long they could stand on their head, I'm still hands-off.

Administrators understood they were not supposed to run the program or determine its direction. Several acknowledged that this stance required self-restraint and was different from how principals traditionally operate. One also questioned whether there was a missed opportunity for alignment between TLI and school/district goals, and was beginning to explore this issue:

I've kind of let TLI do its thing, if you will. Because I feel like any question that the teachers pick as a student learning question is going to be important to them. But now I do want to pull in that alignment so that everything connects.

Several interviewees cited the potential danger of administrators becoming *too* involved with or supportive of TLI, but no one felt this risk of “takeover” had fully materialized.

Effective implementation of TLI required that the administrators *not* be 100% “hands off” but rather that they achieve a delicate balance: give teachers space and independence to direct TLI, but also support their efforts both materially (e.g., designate time, allocate resources including but not limited to funding and coverage) and conceptually (e.g., support new instructional ideas, promote teacher leadership, accept failures).

Administrators also supported TLI by buffering teachers from other distractions:

And trying to figure out how to get rid of all the extra fluff that's going to come so that they can do [TLI]; so it's my job to fight the district when they want to shove something down our throats – no, we're doing this.

Some administrators declined other professional development obligations so they could focus on TLI. Others discontinued feedback systems that they felt conflicted with the SLRC. At times, school administrators had to place themselves

between the school and the district or state. One described how s/he “*pushed back*” on some district initiatives and was able to do so because “*they trust my judgment.*” This individual further reflected that districts needed to give principals the same kind of trust and latitude that school administrators extended to teachers in the TLI.

They’ve almost need to find the right school leader and then find the TLI people underneath them to really make sure that it’s going to work.

We also heard examples of the opposite, such as administrators who used peer observation systems that, in the views of TLI teachers, represented a missed opportunity for alignment if not an outright conflict.

Some participants also recognized tensions between the philosophy of TLI and other accountability forces at the school or district level, which sometimes mandated what teachers needed to do. One person saw TLI as “*trying to give some of the power and decision-making to the teachers*” which s/he embraced, but reality could intervene: “*But your test scores are X, Y or Z, so we’re going to take away your rights to decide.*” S/he pushed this idea further to state that TLI was a better fit in schools that did not have many other mandates or reforms, and thus more appropriate for higher-performing schools.

Finally, some administrators saw their own leadership benefitting from the program. Again, this was more common when they had attended TLI/PLI activities. They grew in skills such as observing student learning and conducting reflective conversations about instruction. Some participants described how these skills could inform evaluative observation processes and saw this as a positive thing.

I am now able to really pick up all of what the students and kids are saying and doing. There’s a very big difference of someone who has trained in that and someone who isn’t.

What [the principal] has done with walkthroughs for this building is s/he's now looked at it like this, "What are the students doing? What is the task that they're working on and what is the teacher doing in relation to that?" That's changed completely from, "Is your objective up?"

Others described how TLI helped them in other aspects of instructional leadership, including motivating veteran teachers to keep growing, helping “Type A” teachers accept failure and peer support, understanding the deeper aspects of school improvement, and remaining patient with the process. Below, one administrator contrasts traditional school management with the type of “organic” leadership TLI encouraged in her:

This is the work that is going to help you change your building....You can be a great manager, but management is only part of the game. Being organized, great, but there is still all this organic-y stuff that lies in between. I think [TLI] helps me with the organic-y stuff, like the deep stuff, and just managing it.

What kind of stuff is that?

It's having those development relationships, the trust and the risk taking; that's not a management thing. You can't manage that. Granted, you're thinking in your head, "How can I make this happen?" but it's got to kind of grow like a plant. You've got to water it and you've got to help it.

Of course, such responses were on one end of a spectrum – not all administrators thought this deeply about TLI – but they demonstrate the program’s potential impact.

Time

Not surprisingly, time emerged as the major barrier to TLI implementation. It was a challenge in all schools and blamed in situations where progress or scaling had stalled. Several specific facets of this challenge were identified:

Finding time to conduct the SLRC

In order to visit classes, collect data, and reflect on the results, participants either had to give up planning time or obtain coverage. Many credited their administrators with being very supportive by hiring substitutes or taking classes themselves:

I could go down there right now and say, “I really want you to come next block,” and s/he would just do it.

Nevertheless, the decision to leave one’s room was not simple. Some described contextual issues (e.g., student behavior, staff turnover, IEP demands) that affected their willingness or ability to leave their rooms. Others identified problems in their master schedules (e.g., lack of common planning, block scheduling) that made it hard to connect with their SLRC partner(s). Participants who co-taught or had interns had a much easier time leaving the room. Finally, we heard about demands such as childcare and coaching, which made it hard to meet after school.

Bringing TLI participants together

As discussed above, convening a larger TLI group was beneficial, but it was even more difficult than finding time for pairs to meet. To some extent, STL supported this process through TLI professional learning days, which gave the community and school teams sustained time to talk. In a couple of schools, TLI participants also had occasional meeting opportunities, or were paid to come in during the summer. In two cases, a regular PLC had been established, although this was not a miracle cure to time difficulties. (For instance, we heard that a PLC meeting had been cancelled because teachers had not had the opportunity to visit each other’s classrooms as often as hoped).

Leaving school to attend TLI/PLI meetings

This was secondary to the challenges of finding time for conducting the SLRC back at school, but still a barrier. Participants could get the necessary substitute coverage but felt conflicted about being away from their students so often, usually on back-to-back days. One stated that it was a professional risk to leave so much instruction with substitute, considering accountability pressures (e.g., DPAS Component V).

Competing demands

Several others also put TLI time issues in a wider context:

There's a lot going on in schools. There's a lot. There's student questions. You have that. You have your RTI. You have your small-group instruction. You have the testing that runs everything. Or drives everything. [TLI] is just another thing ... I think what happens to educators is, there's just too much... There was just an email from my assistant principal yesterday about some kind of training. I did not know what she was talking about.

There is too much happening in this field right now to be able to make sense of it from one day to the next. I've seen initiatives start and literally just die in the air and I don't know why. It's just confusing.

Addressing time problems

We did hear some strategies for mitigating time issues, but they remained a barrier. It was easier for teachers to schedule the SLRC with colleagues who had different planning periods. Thus, in several schools the scaling plan deliberately included multiple grade levels or departments. Another strategy was to embed TLI into times when school was not moving at full speed (e.g., summer, in-service days, faculty meetings). However, administrators' willingness to devote such time to TLI varied, as

did their attitudes about releasing teachers from other obligations. One arranged for his/her TLI teachers to miss part of a school concert to meet together, while another said (as recounted secondhand by a teacher):

I'm not going to subject our staff's morale and sort of like overall outlook, like if you're part of this group you get special treatment because you don't have to be in meetings.

A couple of participants who were very 'bought in' to TLI talked about trying to reframe program activity as part of doing their job well instead of a separate, additional burden.

Teachers think, "This is one more thing to put on my plate. One more thing to do," when really, it isn't one more thing. Because if you do it right and you do it well, then it's not one more thing. It makes the other things that you're doing easier to do.

I always say to [colleagues], it's not hard. And it doesn't take up much time. It's really not a big deal. And it's always interesting. You can always learn something.

Of course, participating teachers' preferred strategy for addressing time barriers was to designate time for TLI collaboration, which often required creative scheduling and/or administrative support. We heard gratitude when this strategy was used or, more often, the wish that it could be.

When discussing the time issue, several participants identified a broader tension within TLI. They knew the program was designed to be teacher-driven and they wanted that freedom, but they needed administrative engagement to help them find resources and surmount barriers.

We get [told] "That's a great idea" but we really don't get support for it, per se. And I get it. It's probably best if it's hands-off from him/her and it doesn't come as a directive from him/her. And then, we're not looking for that but we are looking for some time.

Thus, this barrier connects to issues of administrative role and leadership, discussed above.

Financial resources

Finally, as teachers and administrators looked to the future of TLI, financial issues often came up. The state's budget tightening was a clear concern. Although we did not ask about this directly, several volunteered comments about the price tag for TLI, with one calling it "*steep*." Some described being asked by superiors about the program's costs compared to its value; a couple expressed their own doubts on this issue. The school that withdrew from TLI felt that the costs were hard to justify, thus they chose to incorporate the work in house. At least one other school was beginning to contemplate a "train the trainer" approach in the future to lower costs.

Some took a position of advocacy for the program; they believed strongly in it, worried about future costs, and wanted to persuade us, or others, of its "return on investment":

Is it cost effective to spend X number of dollars for two people? My answer is yes. Because if we can get teachers engaged in their own learning, the key to education is having kids advocate for their own learning. Well, if we can have adults advocate for their own professional development, that would be kind of the same thing.

Look at the growth that you're getting in your teachers. Because if you can make your teachers better, then that goes out to the students. So if I teach 25 kids this year, and I teach of them better than I did the year before, then that's money well spent.

Notice that the second teacher describes the multiplier effect on students, but not on fellow teachers. One administrator noted that Schools That Lead factored scaling into the cost of participation, the argument being that each "green dot" participant also reached multiple others through the program's snowflake model. It is

not clear that participants fully agree with that perspective. Our data suggest that, for Cohort 1 and 2 at least, those who participate fully in TLI as “green dots” are seen to have a different, and more valuable, experience than their peers.

Conclusions

This section summarizes what has been learned in this study, including aspects of TLI that are working well and others that merit attention. We highlight four points. First, we found that the program is mostly working well for individual participants, helping teachers think differently about instruction and learning. Second, we found a dynamic tension between relationships among a small group of “likeminded people” and the need for the program to broaden its reach. Third, we noticed the crucial role of the school administrator in positioning TLI within the broader school context and influencing resources for it. Finally, we note the challenge of aligning TLI with other school improvement efforts.

We found substantial satisfaction with the TLI model. Most teachers and administrators interviewed felt very positively about this professional learning opportunity and found it a refreshing improvement on most PD. TLI felt different because it centered on empowering teachers to investigate student learning. We heard a clear message that TLI was authentic, that the program modeled engaging learning and professional trust in teachers. On one level it is not surprising to find that a group of motivated pioneers mostly found value in the experience to which they devoted two years. Yet we were struck by the intensity of some participants’ advocacy for TLI, and the specificity of the impact examples they shared. These extremely positive responses were not universal. Some teachers spoke with a sense of wistfulness about TLI, recognizing its value and wishing that they had been able to take better advantage or

implement it more fully. Usually in these cases they blamed contextual factors such as the ones discussed above. A small minority of participants seemed to have needs not fully met by TLI, wanting a program that moved faster or emphasized structures as opposed to internal change processes. Some participants and schools found that TLI required more resources (time, funding, human) than they had, and/or found ways to adapt the program to use fewer resources.

As set forth by the micro-credentials, TLI works at three different levels of scale: **Advancing powerful student learning in my classroom**, **Advancing powerful student learning in peers' classrooms**, and **Scaling powerful student learning in my school**. We found that progress towards the first is well underway in Cohort 1 and 2. The transfer of ideas and activities from TLI to the school setting is generally strong. Almost every participant at least attempted to use the SLRC in their own room; most found it a meaningful experience and could describe its impacts in detail. The inquiry cycle seems to “work” in that it yields specific new insights about student learning. These then lead to small “tweaks” that, according to teachers and some administrators, pay off. Our data also demonstrate broader shifts in TLI teachers’ thinking. Whether or not they can consistently act upon these new ideas, they seem to be developing a new image of their classrooms. Administrators confirm these individual level changes, although they more often focus on leadership development rather than instructional practice.

Outside of one’s own room, TLI’s peer impacts are still emerging and vary by school. This makes sense, given the different patterns of scale and collaboration we uncovered. We were struck by how different the snowflake looked in each setting. We found a tension in some schools, where the TLI team was working together well and

influencing each other but only beginning to reach beyond the team. Of course, one possible critique of this program is that it is self-selecting and focuses on “high flyers.” While acknowledging that participants were already strong teachers and leaders, most administrators felt there was value in moving people (as one put it) “from good to great” and that would in turn influence more and different kinds of peers. Still, for TLI to reach its goals of affecting instruction in peers’ classrooms and eventually in the whole school, it does need to spread beyond the small team of “likeminded people” or bring others into the fold. That process was beginning in most schools and well advanced in a couple, but the issue of scale still attracted complex responses from participants. Recognizing this challenge, STL have increased supports for scaling in Cohort 3, an important step forward.

At the school level, TLI’s progress depends on organizational factors: professional culture, administrative leadership, time and financial resources. We also found that these factors are inter-related and specifically that administrative leaders influence the availability of resources for TLI as well as help position the program within the professional culture of the school. Leaders create room for TLI to grow by giving teachers space and ownership over the process while providing support for the program in myriad ways, including prioritizing it, buffering other demands, and supporting a culture of collaboration and risk taking. Our data strongly support the value of educating and supporting school administrators in the program doing more to help them understand their role. Again, STL have already begun to move towards this suggestion, and plan to offer TLI and PLI concurrently in future cohorts. Specifically regarding time, solutions require prioritizing TLI. While specific approaches may vary by school, without designated time, TLI will only work for teachers willing to give up

their free time, substantially curtailing its scope. Solutions to the financial problem are also complicated, especially given the state's budget outlook. At a minimum, talking with school or district stakeholders about how they see TLI's value proposition, especially as it relates to scale, is recommended.

Finally, we see the need to consider and articulate how TLI fits with other improvement work. As the program deepens and spreads, this will be an important issue for administrators to grapple with and STL to support. Schools have only so much energy and attention and TLI requires substantial resources to implement well. It is important for efforts to move in the same direction. We found examples where such alignment had been created, usually through the strong buy-in of administrative leaders and in schools that were not swamped with other initiatives. For TLI to succeed in a wider variety of environments, its change processes must somehow link to other improvement efforts without compromising their teacher-centered, flexible approach.

We often heard that you “can’t understand TLI until you’ve been in it” and that “this isn’t for everyone.” Our data confirm that the change process encouraged by TLI – individual, team, and school-wide – is highly contextual. It cannot be reduced to a generic set of “best practices.” At the same time, now that the program has completed three years, there is an opportunity to tell the TLI story, distill its lessons, and monitor its progress in more solid ways. Various stakeholders (teachers, school and district administrators) spoke of the need for more clear communication about TLI. We also see opportunities to better track program implementation and results without stifling creativity or individuality. The micro-credentials may be a good strategy in this regard.

Finally, the program can use lessons learned from the past as it continues to work with Cohort 3, launches Cohort 4, and plans for its own future and sustainability.

References

- Delaware Department of Education (2017). Reimagining professional learning innovation grant application. Retrieved May 15, 2017 from: [http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/95/CIPD - HE%20PLGrant%2017-18SY.pdf](http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/95/CIPD-HE%20PLGrant%2017-18SY.pdf)
- Delaware Department of Education (2016). Teacher leader pilot overview. Retrieved May 15, 2017 from: <http://www.doe.k12.de.us/domain/424>
- Learning Forward (2011). Standards for professional development. Retrieved May 15, 2017 from <http://learningforward.org/standards>
- Schools That Lead (n.d.). Funders. Retrieved May 15, 2017 from [http://schoolsthatlead.org/STL- Partners.html](http://schoolsthatlead.org/STL-Partners.html)
- Schools That Lead (n.d.). The teacher leadership initiative. Retrieved May 15, 2017 from <http://schoolsthatlead.org/Teacher-Leadership-Initiative.html>
- SocioCultural Research Consultants, LLC (2016). Dedoose Version **7.0.23**, web application for managing, analyzing, and presenting qualitative and mixed method research data (www.dedoose.com).
- Stevens, D., Wallace, D.D., Frankowski, S. & Mash, M. (2016). The Delaware Teacher Leadership Initiative: Growing and Supporting Educator Leadership. A presentation to the NEA National Leadership Summit, Dallas, TX. Retrieved May 15 2017 from https://www.nea.org/assets/docs/LOP126__DSEA_Leadership_initiative_v2__Deborah_Stevens.pdf
- Teacher Leadership Exploratory Commission (2011). *Teacher leader model standards*. Downloaded October 16 2016 from: http://www.teacherleaderstandards.org/downloads/TLS_Brochure.pdf

Appendix 1
TLI Interview Protocols
Interview protocol for participants in the Teacher Leadership Initiative

Interview Questions
1. Tell me a little bit about your role in this school. a. What is your position? b. How long have you been here? c. During that time have you held any leadership roles?
2. Think back to when you first got involved in TLI. a. Why did you decide to join? b. What were your expectations? c. Did administrators and/or peers encourage you to get involved? Why? d. Had you ever done anything like this before?
3. Think about about a student learning reflection cycle you completed in TLI. (If you did several, choose the most memorable). a. What student learning question(s) did you choose to study, and why? b. How did you go about collecting data related to the question? c. What did the data show? What, if anything, were you able to learn about your practice from this experience? d. What happened next? Did you make any changes in your classroom? Did you see any changes in student learning? Describe.
4. The next few questions have to do with peer collaboration within TLI. (Use organizer) a. I see you collaborated with [name(s)] during TLI. How did you decide whom to invite? Probe existing relationship. b. Describe how you and [name(s)] worked together. Where, when did this occur and what did it entail? c. What was it like to work together in this way? How did your relationship evolve? Probe on trust. d. Regarding your collaboration with [name(s)], what do you remember as the most productive part of it? e. What was the most challenging or difficult part? Explain.
5. Ask this question if graphic organizer shows the participant also collaborated in a colleague(s) class. Let's talk about your experiences helping [name(s)] with his/her/their inquiry. a. Tell me more about what that involved (student learning question, data collection method, etc.). b. What did you learn by being in another teacher's classroom?

c. Did this part of the TLI experience affect your practice? In what ways?
6. Let's think a little bit about other people in the school, starting with school leaders. a. How aware were your school leaders about the work you were doing in TLI? b. Did they do anything to support you/ your colleagues in this work? In what ways? c. Did administrators pose any challenges to your work in TLI? In what ways?
7. Now I'd like to hear how TLI interacted with other aspects of your school context. a. How would you describe the adult culture in your school? In general, how do teachers work with each other in this building? b. Did you have the support and resources you needed for TLI? (e.g., time, space) Where did it come from? c. What challenges did you face, or what additional supports would have helped you?
8. Turning now to colleagues in the school other than [those involved in TLI or named]: a. How aware were they about what you were doing in TLI? b. In what ways did these peers support you? Create barriers to your work? c. Do you think your participation in TLI had any impact on these teachers or on the school as a whole? Explain.
9. As you probably know, TLI talks about "scaling the work" or expanding its reach. How did this happen in your school? (use graphic organizer to guide) a. How/why did [name(s)] get involved with TLI? b. What do you know about their experiences? c. What does TLI look like in your school today?
10. We talked about a specific SLRC. Now let's reflect a little more broadly. a. What insights did you gain into student learning through participating in TLI? b. Think about what your classroom looks like today, compared to a couple of years ago (pre-TLI). In what ways (if any) has student learning changed? c. In what ways (if any) have you changed what you do as an instructor?
11. Wrapping up, what advice would you give to a teacher who was considering joining TLI? What advice would you give to his/her colleagues and administrator?
12. What question(s) should I have asked you but did not? How would you answer?
13. Is there any documentation that would help me better understand the impact of TLI on your teaching or your students' learning? (e.g., artifacts or data related to a student reflection learning cycle, information shared at TLI, lesson plans, etc.) If so, would you be willing to share it?

Interview protocol for administrators in the Teacher Leadership Initiative

Interview Questions
1. How did your school first get involved with Schools that Lead/the Teacher Leadership Initiative? a. Whose idea/initiative was it? Who approached whom? b. What were your expectations? What value did you think it would offer to (i) participants (ii) their peers or (iii) the school as a whole? c. How well were those expectations met, and why? (May save this question as a wrap up)
2. Let's talk a little bit about your school context. a. Number of teachers, students b. How would you describe the adult culture in this building? How do teachers generally work together? c. How would you describe your instructional leadership style? d. How does STL/TLI fit with other PD that your school was or is involved with?
3. Who within your school has been involved in TLI? e. Have you personally attended Schools that Lead PD, either geared towards administrators or for teachers? If so: please describe briefly what that PD entailed. (Purpose is to obtain a sense of their familiarity with the TLI process, not to get detailed feedback on their own PD experience) f. Confirm understanding of who was involved with TLI in the school (Cohort 1, Cohort 2, Cohort 3). g. Tell me a little more about that person/those people's role in the school. (Probe on leadership, influence). h. Probe on principal's role, if any, in selecting/ encouraging participation or scaling the project.
4. Let's talk about what you saw happening in this PD experience, beginning with the first participants. (Repeat this question for as necessary for different individuals) f. What is your understanding of what [name(s)] did in TLI? (Probe on understanding of student learning reflection cycle) g. In what ways did [name(s)] collaborate with other teachers during TLI? What did you observe or hear about the nature of that collaboration? h. What changes in [name(s)]'s classroom did you observe, if any? (Probe both on teacher actions and student actions, ask for examples)
5. What was your role as a principal with regard to TLI? d. How did you see your role with regard to this PD?

e. In what ways did you support participants? (e.g., providing access to space/time, supporting inquiry, modeling collaboration, positive pressure) f. What types of successes did you observe with this PD? g. What types of challenges did participants face? h. Did <i>you</i> have the support you needed as a school leader? What additional supports would have helped?
6. We've talked about the experience of individual teachers with TLI. Now I'd like to reflect more broadly. d. Did [name(s)] experiences in TLI have any ripple effects beyond his/her/their classroom? In what ways? e. Describe any impacts you have seen on the adult culture in your school. f. Describe any impacts you have seen on classroom instruction in your school. g. Describe any impacts you have seen on student learning/achievement in your school.
7. Looking ahead d. What does TLI look like in your school today? If your teachers are still participating, what would you like to see in the future? If your teachers are no longer participating, describe this decision. e. What advice would you give to another administrator who was thinking about getting his/her school or teachers involved? Explain. f. What questions should I have asked you that I did not? How would you answer it?

Appendix C

ARTIFACT C: ANNOTATED BIBLIOGRAPHY

Introduction

The purpose of this artifact is to bridge theory, research and practice related to professional learning for educators. This annotated bibliography locates, summarizes, and distills implications from ten frameworks or models of teacher learning. It demonstrates one contribution that universities can make for professional learning practitioners within K-12 schools: synthesizing lessons and ideas from the literature.

I decided to pursue this inquiry for two reasons. First, I wanted to develop my own understanding about how teachers learn and how teacher learning may be connected to changes in instruction and student learning. Doctoral work provided me with an opportunity to dig deeper myself. I hoped that by doing so I would in turn become a more nuanced and skillful professional learning leader (i.e., evaluator, designer, instructor). In a sense, I myself was one “target audience” for this work.

Second and related, I observed both in myself and others the dangers of over-simplification when it came to professional learning. There is a balance to be struck between being “user friendly” and consistent versus over-simplified or one-dimensional. Specifically, I noticed that the Guskey (2000) framework (described in detail elsewhere in this ELP) for evaluating professional learning was in pervasive use throughout Delaware public education. For example, the Delaware Department of Education required it to be the foundation of evaluation plans in competitive grants. It was likewise the basis of almost all my evaluation plans at the Delaware Education

Research & Development Center (DERDC). Surely, I thought, there must be other models out there – including newer or more sophisticated ones?

Also, I was hearing concerns and questions about the Guskey framework raised from various directions including educational leaders in Delaware with whom I discussed my work. They observed that sometimes the five-level framework did not seem to hold; expected changes in teaching and learning did not occur. Why, one asked, if Guskey's Levels 1, 2, 3 and 4 had been attained, was student achievement so hard to budge? Another wanted to understand and “unpack” Level 3 (Organizational Support and Change) more fully since so many professional learning initiatives seemed to get derailed at that level. Others wondered, are the changes always so linear? Do they always happen in that order? Do different kinds of educators/participants experience “levels of impact” in different ways? I wanted to offer explanations to such audiences or at least alternative ways of conceptualizing teacher learning – and, therefore, designing, implementing and/or researching and evaluating programs to encourage such learning. For these reasons, I anchor this artifact in comparisons to Guskey's (2000) work. At times, I also compare or contrast the models I review, but this is not a fully synthesized literature review.

I used various search techniques to locate these frameworks. I started by revisiting a few articles I had already read, specifically the more conceptual or theoretical ones (e.g. DeSimone, 2009). Since I intended to draw comparisons to Guskey's framework, it made sense to read further into his work (e.g., Guskey, 2002). I then searched the reference lists in these articles and looked for subsequent articles that cited them. Searching in library databases I discovered journals that I had not known about, especially *Professional Development in Education*, which is based in

Great Britain and publishes research from many international contexts. Coldwell and Simkins (2011), Kennedy (2014) and King (2014) all came from this journal. I found that looking outside of the United States (even at other Anglophone countries) brought a valuable new perspective. Following this international line of inquiry led me also to Timperley (2011) and Evans (2014). In efforts to read a wide swathe of research I searched for literature reviews (e.g., Opfer and Pedder (2011), and Kennedy (2016)).

Late in my research, I discovered a new article (Boylan, Coldwell, Maxwell & Jordan, 2018) that examines five of the frameworks I had previously identified (i.e., listed alphabetically, Clark & Hollingsworth (2002), DeSimone (2009), Evans (2014), Guskey (2002) & Opfer & Pedder (2011)). The article analyzes five aspects of the models: components, scope, theory of learning, location of agency, and philosophical paradigms. I approached this article as confirmation that I had located relevant and appropriate models. I analyzed the Boylan et al. article as a secondary source to help me understand the five frameworks it discusses. One specific contribution of that article is to locate and describe ways each model has been used in subsequent research.

Two points need explanation. The first relates to vocabulary of “research” and “evaluation.” While the broad purpose of research is to study phenomena systematically to generate new knowledge, in evaluation it is more applied. Especially in program evaluation, the goal is to generate recommendations that can improve programs (in this case, for professional learning). Some but not all of the authors here distinguish whether their work applies to “research” or “evaluation.” Second, this bibliography is clearly not exhaustive. I focused on works that explicitly aim to model, or represent, how teachers learn. Many of the models connect professional learning activities to changes for both teachers (in knowledge, skills, beliefs and/or practices)

and/or students (in achievement or other outcomes). Most include graphical representations. Most articles also discuss one or more empirical studies, either conducted or reviewed by the author(s), upon which the model was built. I acknowledge there are other works that could be considered “models” but are not included in this bibliography (e.g., Borko (2004), Mouza (2009)).

There are multiple potential audiences for this artifact. Since I review ten different frameworks, some lend themselves better to research and others to practice; some also pertain to policy. Some authors include explicit statements of their frameworks’ implications or potential uses. In other cases, I inferred and identified these applications myself. Each write-up identifies applications by audience: PD designers, implementers, supervisors, evaluators and/or policymakers. Ultimately this artifact fits within my ELP’s theory of change since it aims to produce better-informed educational practitioners *and* stronger PD evaluations.

1. Guskey (2002)

Summary

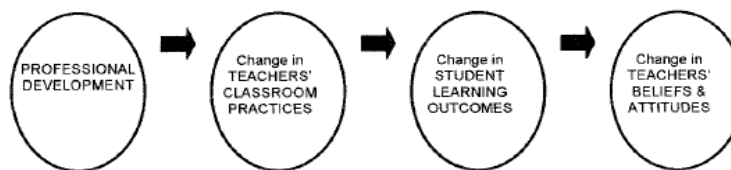


FIG. 1. A model of teacher change.

Figure C1 Guskey (2002) model

Guskey posits a linear model of teacher change. As shown in Figure 1, professional learning experiences lead to change in teachers' classroom practices, then to changes in student learning outcomes. When teachers see changes in their students (i.e., identify that new practices are effective), they shift their beliefs and attitudes. For Guskey, the order of events is crucial and he presents his model as an alternative to the more common view that changes in attitude predate and are necessary for changes in practice. He argues instead that "change is primarily an experientially based learning process for teachers" (p. 384). The article includes examples of research studies supporting this model, but it is not a literature review (e.g., no criteria are provided for how articles were identified). No discrepant examples are provided.

How does this framework add to or contrast with Guskey (2000)?

I am not sure how Guskey himself would account for the difference between his two models. He based his 2002 framework on earlier writings so the chronology is not clear. I see them not as contradictory but as different in purpose and scope. The model of teacher change (2002) has narrower boundaries than the five-level program evaluation framework (2000). It focuses on change at the individual teacher level. The 2000 framework is used to evaluate PD programs more broadly. It includes components not represented in the 2002 model, such as organizational change and student learning outcomes. To think about this another way, look at the first arrow in the 2002 model of teacher change, which connects PD to changes in teachers' classroom practices. In the program evaluation framework (2000), such changes occur at Level 4. Before that, PD must satisfy and meet teachers' needs (Level 1). Teachers must gain new skills/knowledge (Level 2) and receive organizational support (Level 3). So putting the two models together results in complications. One critique of both

Guskey models is that they are too linear. Guskey acknowledges that “exceptions to this model certainly exist” and specifically that beliefs must shift at least “from ‘cynical’ to ‘skeptical’ for any change in practice to occur” (p. 385).

Who should know about this framework? How can they apply it?

Although published in an academic journal, Guskey highlights implications for practice especially for those who plan or oversee professional development programs. These include:

- Recognize that change is a gradual and difficult process for teachers

- Ensure that teachers receive regular feedback on student learning progress

- Provide continued follow-up, support and pressure (p. 386-8)

The first two implications both connect to the idea that teachers are motivated by student learning. Teachers may be concerned with trying new things or risking failure at their students’ expense.

Timely formative data on student learning is essential to help move the wheels of change. The model has implications for what activities should be emphasized in professional learning programs (i.e., more time spent on mastering the practice and obtaining formative feedback about it; less time persuading teachers to “buy in”).

Implication (2) provides support for professional learning that includes or is focused on detailed evidence of student learning. Implication (3) is important for school leaders/supervisors as well as coaches, facilitators, etc. as they develop plans to follow-up on professional learning and hold teachers accountable. This model relies on teachers beginning to implement new practices as an early step in the theory of

change. Ensuring that this happens, and responding to teachers when it does, or does not, is important for successful implementation.

Guskey also identifies several areas of his model that merit more research. These include the best ways to provide feedback about student progress to teachers and to help teachers move towards instructional action. Guskey does not explicitly indicate implications of his model for research or evaluation design. Nevertheless, this model as well as Guskey's five-level framework (2000) are used as the foundation of many research studies and evaluations. Boylan and colleagues (2018) reference several examples from disciplines including physical education and science. Guskey's model suggests that researchers should look for changes in practice before changes in beliefs. It might affect the way that evaluators conceive of programs (e.g., the logic model).

2. Desimone 2009

Summary

This piece is targeted at researchers and aims to improve the quality of studies related to professional development (PD). It does this by elaborating what "counts as PD" (p. 182), identifying which features of PD have empirical evidence, and then to propose a "core conceptual framework" of teacher learning and to discuss its implications for research. In part 1, Desimone argues there is a research consensus of features that make PD effective: content focus, active learning, coherence, duration, and collective participation. In part 2, she presents a theory of action for how PD leads to changes in student learning (Figure 1). This is presented as a linear, causal path. However, she states that relationships between PD features, teacher learning, instructional change and student outcomes can be "interactive and non-recursive," as

shown by the dual arrows in Figure B2. Like Boylan et al. (2018) I found this to be a bit self-contradictory.

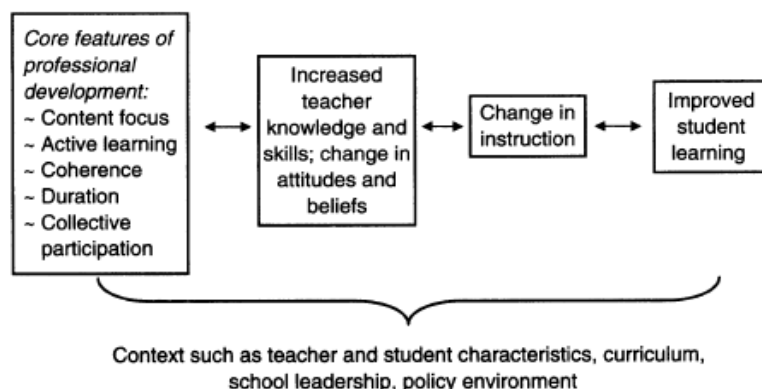


FIGURE 1. *Proposed core conceptual framework for studying the effects of professional development on teachers and students.*

Figure C2 Desimone (2009) model

How does this framework add to or contrast with Guskey (2000)?

Like Guskey’s, Desimone’s model follows a linear path beginning with the professional development activity and ending with student learning outcomes. As Boylan et al. state (2017), both models view the stimulus for change as an external PD activity or event. Desimone analyzes those PD activities in more detail, and indeed the “core features” seem to be the most often referenced part of this article. As shown in Figure C2, DeSimone treats context (e.g., “teacher and student characteristics, curriculum, school leadership, policy environment”) as an overarching mediator and moderator for any and all parts of the pathway. This contrasts with Guskey (2000) in which such issues are explicitly but discretely located within the model (i.e., at Level

3). Finally, Desimone arranges the steps of the change process in a different order than Guskey (2002). Teachers' knowledge, skills and/or beliefs change before their instructional practice. Attitude change can be caused by professional learning experiences, and "change in belief leads to change in practice leads to change in students" (p. 395). Desimone is not clear whether teachers must experience change in all three areas (knowledge, skills and beliefs) but states that the teacher's personal changes must occur before they try new things in the classroom.

Who should know about this framework? How can they apply it?

Desimone writes to persuade about the value of her framework. She seems to be saying to fellow researchers: don't make it so complicated, we can all agree. She argues that a more consistent approach to studying PD would be more useful for practitioners, allowing better planning and oversight of PD, as well as advance research and understanding in the field. In the final section of the article, Desimone discusses ways to empirically test her framework. She specifically focuses on pros and cons of different data collection methods for "measuring professional development and its effects on changing teacher practice" (p. 188). This section has implications for research design and instrumentation. She disputes what she calls a "bias" that observations and interviews yield more valid data than surveys. She argues that surveys can work well if they ask about specific, concrete changes in practice rather than general self-reports. Finally, Desimone ends with questions and ideas for further study, including the potential of new data collection methods (e.g., vignettes), questions about impacts of PD on non-volunteers, questions about thresholds (how

much PD is enough?) and about the key content knowledge for teaching (i.e., pedagogical content knowledge).

While this piece clearly has a research audience, Desimone has also written for practitioner journals (2011). The “core features” are particularly practical. They can drive the design of professional learning activities and can be used as criteria for the types of activities most likely to be effective. For example, a district or school leader could decide whether or not to support a professional learning activity based on the extent to which it demonstrates these features. Policymakers could use them in similar ways. However, I note that later researchers (e.g., Kennedy (2016), Opfer & Pedder (2011)) warn against applying the features reductively. It is not simply a question of whether or not they are present, but in what ways, at what intensity/quality, for whom, etc. Finally, this model can be used to structure evaluation of professional learning activities, even internally (e.g., by a school or district leader). Like Guskey, Desimone argues that even busy practitioners need to dig deeper than teacher satisfaction and measure (1) change in skills/knowledge/attitudes (2) change in instruction and (3) change in student outcomes.

3. Clarke and Hollingsworth (2002)

Summary

The authors describe their Interconnected Model of Professional Growth (IMPG) developed based on empirical studies, mostly involving math instruction/PD in Australia. This article provides citations for all the studies and selected evidence, mostly qualitative or case study, to illustrate its points. It is based in a view of “teacher change” as a personalized learning process that can occur in multiple ways. This

article begins with a review of other models including Guskey (2002). The authors argue that the IMPG, illustrated in Figure B3, can incorporate all previous linear models.

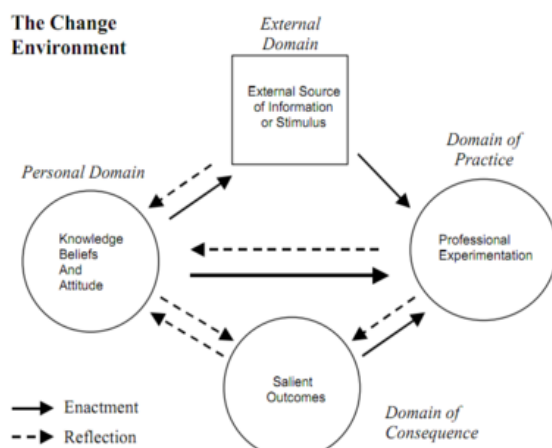


Fig. 3. The interconnected model of professional growth.

Figure C3 Clarke and Hollingsworth (2002) model

The IMPG includes four domains. Only one is external to the teacher. The “source of information or stimulus” could be a PD activity but it could also be less formal (e.g., interaction with another teacher). The other domains encompass the teacher’s knowledge, beliefs and attitudes (Personal), experimentation including, but not only, in the classroom (Practice), and outcomes s/he views as salient (Consequence). Change may occur in any one or more of these domains, in any order. The model encompasses many potential patterns, some of which are illustrated in the article. Solid lines represent enactment (consciously doing something new) and dotted

lines show reflection. The authors distinguish between “change,” which can be a fleeting moment of experimentation, and longer lasting “growth.” They also emphasize teacher inference. Two teachers may interpret the same event differently. For instance, increased student talk could be interpreted as an outcome of increased engagement or loss of teacher control. What matters for understanding teacher change is the interpretation of the events that occur: “the practices of the classroom are co-constructed through the actions and the inferences of the participants” (p. 956). The model is constructivist.

The final section describes how school context affects “access, participation, experimentation and application” of PD. Examples of facilitating or constraining factors include tangible or human resources, pedagogical ethos, professional culture, or collective participation.

How does this framework add to or contrast with Guskey (2000)?

The authors intentionally build on prior models including Guskey and what they call the “implicit model of teacher change” (similar to Desimone, 2009). While these are both single pathway models, the IMPG has multiple pathways and the article shows several different ways change can happen. Although both Guskey and Desimone *acknowledge* some variation, the linearity of their models comes across more strongly. The IMPG is inherently more flexible. This is presented as a strength: just like students, teachers learn in different ways: “Unlike more prescriptive models, the alternate pathways in the interconnected model allow us to give recognition to the idiosyncratic and individual nature of teacher professional growth” (p. 965).

This model also distinguishes between “enactment” and “reflection” and in general makes more room for teacher sense making. For example, a key assumption in

Guskey is that the ultimate outcomes of professional learning occur for students (primarily, in achievement or learning gains). The IMPG has room for a variety of outcomes, including but not limited to student outcomes, and teachers determine which are salient for them. As Boylan et al. note (2017), teachers have more agency in this model than in some of the others.

Who should know about this framework? How can they apply it?

This article appears written for a research audience and is widely cited in subsequent studies, literature reviews, etc. (Boylan et al., 2017). The authors identify analytical, predictive, and interrogatory applications. The model can be used to analyze data, such as by change domain, individually or in patterned combinations (what the article calls change sequences or growth networks). It can also generate *potential* or predicted patterns, which could then be tested empirically. Finally the model can be used to frame research or evaluation questions and designs.

Although not fully discussed in the article, Clarke and Hollingsworth's model also has implications for practitioners. Its "interrogatory" function could also be used by designers. The idea that teacher growth occurs in multiple ways and varies by individual is important for professional learning designers, facilitators, and supervisors to keep in mind. The model also reminds us that growth can occur independent of any external stimulus. While teachers do care about student learning and achievement, this model reminds us that other outcomes (e.g., feeling of mastery, work/life balance) might also be important and motivating for teachers. Finally, it places a lot of emphasis on reflection as a mechanism for change and growth. This suggests the importance of including reflection opportunities during and after professional learning activities.

4. Opfer and Pedder (2011)

Summary

This is a literature review. It argues for a conceptualization of teacher learning informed by complexity theory and states this is necessary because correlational research has so often been disappointing; a PD activity can have all the research-based features and still not bring about change in teacher practice. The authors argue it is essential to dig deeper into *why* teachers change their practices; only research that explicitly engages this question was included in the review. Three main shifts in conceptualizing teacher learning are proposed. We must recognize (1) context and organizational conditions i.e. some things work in some places, for some people (2) the ‘Goldilocks’ principle and intensity i.e. too much of a good thing can hurt, for instance too much collaboration can lead to group think and (3) there are nested systems – individual teacher, organization/school, and PD activity – that interact to influence learning.

The review then considers each system. For both the individual and the school organization, the point is made that a certain amount of dissonance foments change (“complex systems need to be off balance to move forward,” (p. 393)). There needs to be a balance between internal and external stimulus for change. Conclusion: “To understand and explain why and how teachers learn, we must consider how a teacher’s individual learning orientation system interacts with the school’s learning orientation system and how both of these systems together affect the activities (and features of activities) in which teachers participate and then are reciprocally affected by the changes that occur from participation in those activities” (p. 393-394). Perhaps fitting for a model characterized by complexity, there is no visual summary.

How does this framework add to or contrast with Guskey (2000)?

Opfer and Pedder critique pathway or level models such as Guskey (2000, 2002) and Desimone (2009) as simplistic. They argue that the change process is non-linear and can begin with teacher attitudes, practices, or student outcomes, and, further, that “change in only one area may not constitute teacher learning” (p. 396). Although the article does not explicitly define teacher learning, it seems to encompass changes in beliefs and in practices, which result in changes for students. It is a ‘complex system, rather than...an event’ (p. 378)

Another contribution of this framework is its emphasis on individual teacher orientations. Similar to Clarke and Hollingsworth, this model creates space for individual differences and preferences among teachers. “Individual orientation to learning system” is defined as “the interaction and intersection of knowledge, beliefs, practices, and experiences” (p. 388).

Who should know about this framework? How can they apply it?

The article’s main purpose is theoretical (Boylan et al. (2017)). It proposes implications for research, overall that it needs to better attend to complexity and context: “we need more studies that investigate how the generative mechanisms of teacher learning appear in different combinations and sequences, with different weights, in different but concrete situations” (p. 394). They call for “complexity-influenced research designs” (p. 396), which includes attention to multiple systems and a balance between contextual specificity and generalizability. However, the article is not very concrete about what this looks like.

Although not intended for practitioners, some ideas in this article may be useful. It is more a question of using insights from the article to inform thinking and

planning than applying the entire framework. For example, the point that both individuals and school systems benefit from some disequilibrium may inform professional learning leaders. Striking a balance between internal and external stimuli for change is important in planning professional learning. The idea that we must consider intensity and context, not just the presence or absence of certain program features, also could guide design and implementation.

5. Evans (2014)

Summary

Focused on the UK context, this article has two purposes: to propose a theory of teacher professional learning and then discuss its implications for educational leaders. Evans purposely wants her model to be “context free,” universally applicable, and theoretical. It is focused on the micro level, that is, the individual teacher’s cognition including what she calls “private epiphanies” (p. 186). The model includes only the teacher and his/her professional practice; students do not appear. According to Evans, while in the US research in this area focuses on improving student outcomes, “in other geo-cultural contexts, the professional development of teachers is considered a justifiable end in itself – a worthy focus of study, irrespective of whether or not it may be seen to lead to gains in relation to pupil learning” (p. 181). In Evans’ model, the key driver is a teacher’s recognition that there is a “better way” of doing things or solving instructional problems. This recognition drives change and the “transfer of loyalties” to new practices (p 187). This can occur in informal (and even unconscious) ways as well as in response to formal learning activities.

The model is based on a three-part conception of a profession (and thus, PD). In pursuit of a “better way,” teachers may change behaviors, attitudes, and/or skills and knowledge. Evans distinguishes between professional development, which may only affect behavior, and professional *learning*, which also needs to touch the intellectual and/or attitudinal domains.

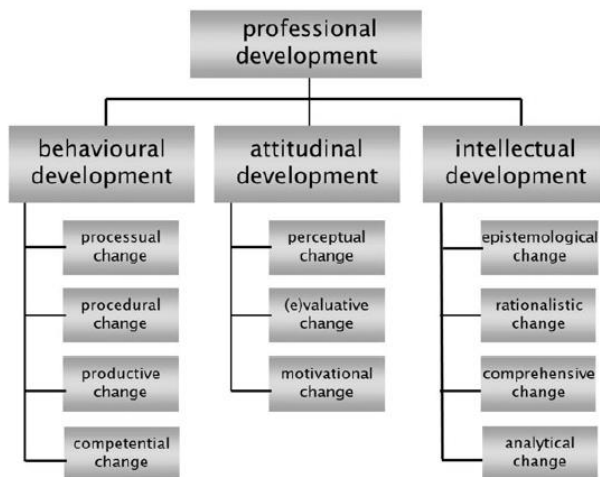


Figure 2. The componential structure of professional development.

Figure C4 Evans (2014) model

How does this framework add to or contrast with Guskey (2000)?

Evans discusses the Desimone and Guskey models, describing their value but also limitations for understanding why teachers develop. For instance, she questions Desimone’s core features: “the bases of their efficacy and potency remain

unexamined. What this means is that this conceptual framework offers no reliable capacity for explaining...deviance, atypicality, relationality, and causality” (p. 183). Evans prefers Clarke and Hollingsworth’s model but takes issue with some of its assertions, such as that the only two mediators are enactment and reflection.

Who should know about this framework? How can they apply it?

Evans’ primary audience is school leaders. She argues that they need to understand professional learning to lead it effectively. Although intended to offer lessons for these practitioners, the article is dense and theoretical. A key takeaway is the distinction between behavior change based in compliance and learning which also affects the other domains (attitudes, knowledge).

PD – like the professionalism that it is intended to enhance – is multidimensional; it is not simply or narrowly about changing people’s behavior – how they do or go about things, or how much they do or produce, or what generative effect their changed practice has – it is also about changes to their attitudes, intellectual capacity and mindset (p. 193)

If school leaders want to help teachers move in a particular way, they need to find ways to help teachers view new practices as a “better way”: “The importance of winning over hearts and minds cannot be over-emphasized” (p. 195).

6. Coldwell and Simkins (2011)

Summary and comparison to Guskey (2000)

The purpose of this article is to critique “level models” of professional development especially Guskey (2000) because they: (1) rest on flawed assumptions that higher levels are most useful and that each level follows the other and is caused by it; (2) misunderstand Level 3. It is “not a *consequence* of the previous stage as the

other levels are but as a set of *conditions* for the previous stages to lead to the next ones” (p. 145); and (3) do not sufficiently explain why changes occur, or do not, especially in student learning. The first two critiques are unique to this article; the third echoes other articles reviewed here.

The authors propose a model with many additional components. The moderating factors are similar to Guskey’s Level 3, but they are stated more explicitly and pulled out of the “pathway” to instead be an overarching influence. The antecedents put more emphasis on an individual teacher’s expectations and motivation.

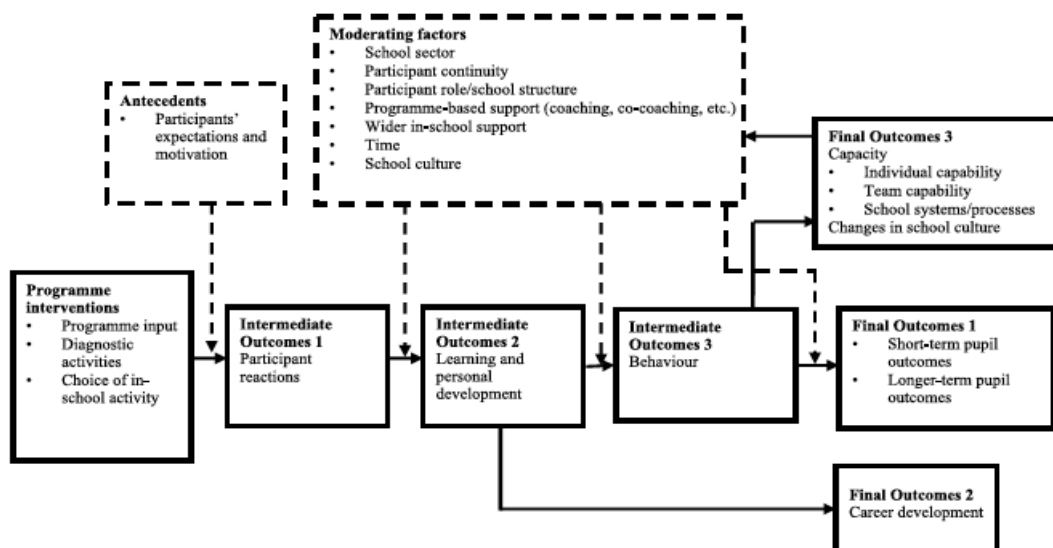


Figure 1. A basic model of leadership programme effects.

Figure C5 Coldwell & Simkins (2011) model

Who should know about this framework? How can they apply it?

This article is targeted for evaluators. The final section contrasts three ontologies for professional learning evaluation: positivist, post-positivist, and

constructivist. Positivist evaluators typically use experimental or quasi-experimental methods and focus on causation. Post-positivists focus on “uncover[in]... combinations of context, mechanisms, and outcomes. These approaches have a strong focus on learning from evaluation about why and how programmes work, not just “what works” (p. 151). Constructivists believe “that programme purposes may be contested, that individuals may experience interventions in different ways, and that understanding these contestations and experiences may provide important information that can contribute to our understanding of how interventions work” (p. 152).

Level models of PD are incompatible with constructivist approaches because they are “instrumentalist” and assume a clear-cut set of objectives, defined externally (p. 153). The authors acknowledge the accountability purpose of evaluation, especially of publicly funded programs, but want to move beyond “Did it work?” or even “How did it work?” to ask deeper questions, such as “For whom does it work?” or “What is meant by ‘work’?” Questions like these would affect the design of research or evaluation studies, as would the ontological position more broadly. However, the article is not always concrete in showing how to use these questions in the context of a program evaluation that will be useful to practitioners. One fairly solid point is that in ongoing professional learning (what in the UK is called “continuing professional development,” or CPD) the boundaries of the program/activity/evaluand are not always clear. What is “in” the program and what is “outside” of it can be complicated to tease out. Moreover, different stakeholders may have different views of this issue. In general, the article encourages understanding professional learning through the eyes of the person experiencing it – and this has implications for both research and leadership.

7. King (2014)

Summary

The article describes the drafting, testing and revision of a framework for evaluating the impacts of teacher professional learning activities. The context is an Irish multiple case study examining long-term impacts of a literacy initiative in five urban schools. One difference between this framework and others is its explicit focus on sustainability of impacts. King also summarizes many prior frameworks for evaluating training/PD including Kirkpatrick (1959), Guskey (2000, 2002) and Bubb and Earley (2010), which built on Guskey. It also incorporates the Levels of Use from the Concerns-Based Adoption Model (CBAM) (Hall and Hord, 1987). The study involved matching research questions to a provisional framework for evaluating PD, then updating it based on the results. The final framework is overleaf.

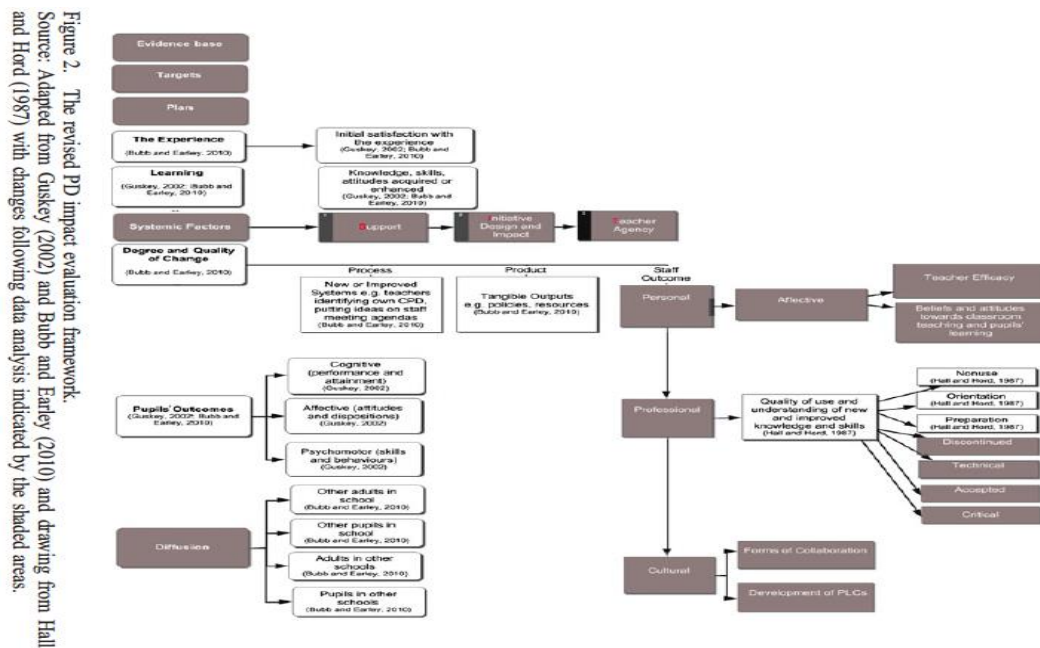


Figure C6 King (2014) model

How does this framework add to or contrast with Guskey (2000)?

King's model combines parts from Guskey as well as previous models (cited in the figure above) with additions from her data. She argues that her model improves on prior models by (1) addressing systemic factors more completely (2) incorporating a wider range of staff outcomes (personal, professional and cultural) and (3) considering diffusion to educators and students in the school, or beyond it.

King describes three groups of systemic factors. The first overlaps with Guskey's Level 3: support, particularly from leaders. King found that "leadership support was the mechanism through which other supports, such as the development of PLCs and the modeling of practices by an advocate (someone who is driving and supporting the practice) were enabled to develop" (p. 102). The other systemic factors are less explicit in Guskey, so they represent a contribution. King found that "initiative design and impact" was a supportive factor (i.e., one reason the initiative had such lasting impacts was its research-aligned design). This matches up more closely with Desimone's (2009) "core features." Finally, King identifies teacher agency as another systemic factor. Like some of the other more recent models, she views the Guskey model as not saying enough about agency.

Related to this idea, King also expands upon the impacts of PD on teachers (i.e., Guskey's Levels 2 and 4). Regarding professional outcomes, she looks at the *quality and extent* of new practices, not just whether or not they are used. She puts more emphasis on personal outcomes (i.e., beliefs/attitudes, efficacy) and cultural outcomes (i.e., increased collaboration with others, including in PLCs). An important idea in this framework but not present in Guskey is that, over time, professional learning affects the organizational climate for adults and students within a school and, eventually, beyond it. This is represented in the Diffusion section.

Who should know about this framework? How can they apply it?

King's additions result in a comprehensive but hard to use model. It illustrates trade-offs in modeling teacher learning. King states her model can be used by leaders or teachers but does not show how. One possible application lies in the range of teacher outcomes. This encourages us to think more broadly about what teachers "get out" of the professional learning experience, including how it affects attitudes, beliefs and efficacy as well as how they work with others in or beyond the school. This insight could be used for practice (e.g., for self-assessment in the teacher evaluation system) or research (e.g., by developing items for data collection instruments). Another takeaway is the importance of systemic factors. In a later article, King (2016) digs further into this and suggests a range of questions related to support, design/impact, and agency, which could be used for planning, implementing, or evaluating PD programs.

8. Kennedy (2016)

Summary

This is a new literature review. It argues that a more nuanced understanding of what "teaching" entails is important for answering the title question. Kennedy thus uses two questions to structure the review: *What problems of practice do programs aim to inform?* (p. 946) and *What pedagogy do programs use to facilitate enactment of their ideas?* (p. 947). Each is broken into four categories, listed below, which can be combined in various ways. According to Kennedy, problems of practice include to: (1) portray curriculum content (2) contain student behavior (3) enlist student participation and (4) expose student thinking. Program pedagogies include: (1) prescription (telling teachers what to do) (2) strategies (communicate a goal and practices to reach it) (3)

insight (raise questions, encourage reflection) or (4) body of knowledge (share information, let teacher decide what to do with it).

The review mixes and matches these categories to organize its findings. For instance, it looks at studies of prescription related to curriculum content. Kennedy used strict criteria for inclusion: only experimental studies published in the United States since 1975 that include evidence of student learning, span a year or more, and control for participant motivation to learn (e.g., whether or not the program was voluntary). Kennedy found a major difference in effect size on student achievement outcomes for programs with volunteers (0.16) versus non (0.03).

How does this framework add to or contrast with Guskey (2000)?

There is no inherent contradiction between Kennedy's review and Guskey's five-level framework. Kennedy offers a more refined way to categorize and understand professional learning programs. She adds detail to the "front end" of Guskey's model, which might help explain subsequent outcomes.

The stronger contrast is with Desimone (2009) and other frameworks that attempt to itemize "features" of effective professional learning programs. Many of the studies in Kennedy's review claimed to be aligned with these features, yet not all were found to be effective on student learning. Kennedy provides several explanations why this is so. First, programs focused *exclusively* on content knowledge were less effective than those that also examined student learning. Second, while both PLCs/collective participation and coaching are said to be "research-supported," Kennedy found a range in quality of these features. It is not enough to *have* a PLC or a coach; it has to be an effective one. Related, Kennedy emphasizes the quality of professional learning facilitation and states this is under-explored in other frameworks.

Finally, she makes the common sense point that duration and intensity is not in and of itself beneficial. Prescriptive approaches are generally not effective, so more would not be better.

Who should know about this framework? How can they apply it?

Kennedy is writing for a research audience. Similar to other authors discussed here, she calls for a deeper understanding of teacher learning to improve the quality of research in this area. She points out that much less is known about teacher learning versus student learning, and what is known does not sufficiently influence practice. She identifies specific areas for further study, for instance the role of motivation in learning and the “slow and incremental way in which teachers incorporate new ideas into their ongoing practices” (p. 29). This second point demands longer-term research.

Although not written for practitioners, this article has plenty of implications for them. Kennedy highlights the “disjuncture” between learning science and professional learning design. For instance, “programs such as the Los Angeles Science Immersion program which aims to actively immerse students in scientific activities but at the same time inundates teachers with volumes of prescriptive details about how they should immerse their students in science. Why would we expect these detailed prescriptions to work for teachers if we do not believe that they work for students?” (p. 973). She offers more specific considerations for designers, too. For example, given her findings about the impact of facilitation quality, we might raise questions about “train the trainer” approaches to professional learning, or at least insist on a high bar. Similarly, those planning programs should attend to the quality of PLCs and coaching (using research in each area for guidance), and those implementing or evaluating programs should monitor and measure it. Kennedy asserts that in programs,

“attendance [may be] mandatory, but learning is not” (p. 973). Her findings may encourage leaders to lean towards letting teachers choose professional learning opportunities, or at least working harder to obtain their engagement. Finally, the observation that providing teachers with content knowledge is not sufficient has implications for professional learning design. It suggests that programs should include examination of student learning. This final point seems consistent with the Guskey model.

9. Kennedy (2014)

Summary

Similar to Kennedy (2016), this article provides a way to categorize and understand continuing professional development (CPD) programs. The context is British. This article revisits the author’s 2005 CPD taxonomy in light of research and policy changes. The original framework categorized nine models or types of CPD and three purposes for which the model can be used. The framework was slightly updated in 2014.

Table 1. Spectrum of CPD models (adapted).

Purpose of Model	Examples of models of CPD which may fit within this category
Transmissive	Training models Deficit models Cascade model
Malleable	Award-bearing models Standards-based models Coaching/mentoring models Community of practice models
Transformative	Collaborative professional inquiry models

Figure C7 Kennedy (2014) model

This model distinguishes between three purposes of CPD. The two poles of the spectrum are “transmissive” (i.e., aiming to communicate specific content or practices) and “transformative” (i.e., empowering teachers to collaborate, inquire, and make independent decisions). In the middle are “malleable” CPD models that can serve either purpose. For example, Kennedy claims mentors can help teachers develop their autonomy, or can encourage them to conform. Kennedy pushes one step further, examining different theories of professionalism and contrasting the “managerial” and the “democratic.” These align with “transmissive” and “transformative” purposes above. A managerial perspective on professionals emphasizes compliance, consistency, and accountability. A democratic perspective “privileges collaboration, openness, teacher agency, and an overt commitment to social justice” (p. 695). Kennedy’s article implies a preference for more transformative and democratic

approaches – or at least a sense that in the current political climate, such approaches need to be more valued and restored.

How does this framework add to or contrast with Guskey (2000)?

For Kennedy, CPD is both a pedagogical and a policy construct and much of her article focuses on policy. She argues that CPD is increasingly framed as a tool for economic development and educational reform. Correspondingly, both policy and practice related to CPD has narrowed to focus on measurable student outcomes associated with educational/economic success. Kennedy's political and sociological focus is different from Guskey's but his model does not conflict. Indeed, Kennedy cites Guskey as well as many other models of researching professional learning. Unlike some of the authors discussed here, she does not critique Guskey. In fact, she views his model as holistic in that it acknowledges multiple forms of professional learning impacts: "change in the classroom practices of teachers, in their attitudes and beliefs, and in the learning outcomes of students" (2002, p. 381).

Who should know about this framework? How can they apply it?

Kennedy's main audiences are researchers and policymakers. Like several authors reviewed here, she finds existing research on professional learning "partial," "fragmented," and "under-theori[z]ed" (p. 689). Kennedy wants to develop "sophisticated but accessible means for understanding CPD more deeply" (p. 690). A general framework such as the one proposed here provides a useful analytical tool. Kennedy proposes concrete applications specifically for policymakers at a system level (e.g., state or district). First, the model can help them interrogate the underlying assumptions and purposes of their CPD and ensure that activities align with them.

Second, the framework provides a “shared language for discussion and debate” (p. 695) about CPD policy. Kennedy argues for the importance of understanding context, especially the political context, within which CPD operates. Therefore, she specifically cautions policymakers against using a simplistic approach of borrowing or replicating CPD that might be successful elsewhere. “What works” in one context may not in another; policymakers must attend to *where*, *how* and *why* a CPD program or initiative works.

10. Timperley (2011)

Summary

This book is a combination of theory, literature review, and a report on a 10-year literacy initiative in New Zealand. It includes two frameworks, the teacher inquiry and knowledge-building cycle and a corresponding cycle for leaders. The following are notes about each chapter of the book, which correspond to steps in the teacher cycle.

Finding out about students: Professional development needs to be based on actual student learning needs. Data that teachers trust creates motivation and a “need to know.”

Building teacher knowledge: Learning needs/goals should be grounded in student data, with room for teacher input. Teachers may not know what they don’t know, thus should not have total discretion for selecting their goals/needs.

Checking new opportunities for students: Implementation is part of the learning process for teachers, not an outcome of it. Related, teachers can formatively evaluate and monitor their own implementation of PD.

Role of school leaders: This is a major emphasis in this model. The strongest principals view their teachers as a “class” and hold themselves accountable for each one’s professional learning.

Role of facilitators. PD facilitators also have a “class” of adults whose buy-in and engagement they need to secure. Challenges include resistance, dissonance, “over-assimilation” (p. 142-3) in which teachers think they have changed practices but do so only superficially, and the need to engage leaders. Grounding PD in student learning can help avoid “competing theories about practice” (p. 156) and threatening teachers’ sense of professionalism.

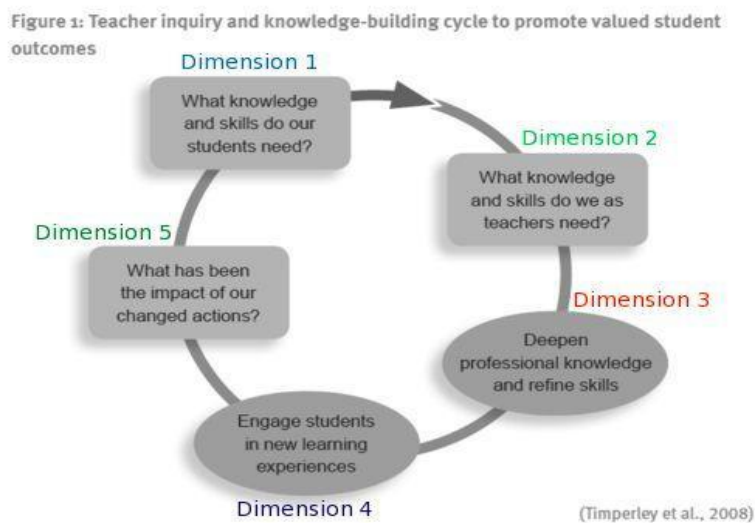


Figure C8 Timperley (2011) model

How does this framework add to or contrast with Guskey (2000)?

Timperley's framework matches up well with Guskey's suggestion to plan backwards from desired student learning outcomes, i.e., to use the evaluation levels in reverse as a guide for planning (2000). Compared to sequential frameworks like Guskey's, however, Timperley integrates learning, implementation, and reflection. Timperley also looks more broadly at professional learning systems and their actors. Alone among the models discussed here, this one discusses implications for professional learning facilitators and *their* own learning. There is also substantial focus on the role of the school leader and on system-wide learning and inquiry. Finally, Timperley is explicit about the importance of participant engagement. Without engagement, even professional learning activities that use research-supported designs and "should" be effective may not be.

Who should know about this framework? How can they apply it?

This book and framework could be of interest to researchers, policymakers or practitioners. It is written in an accessible style with many useful and practical "nuggets." For example, the idea that teachers "may not know what they don't know" suggests that instructional leaders should be involved in the process of selecting professional learning activities. (Though Timperley still thinks teacher choice and ownership is very important, too). Anyone charged with designing or delivering activities for teachers could benefit from Timperley's analysis of those challenges.

Conclusion

This bibliography has summarized ten frameworks for professional learning, all published since Guskey (2000), and distilled implications for practice, research/evaluation, and/or policy. What have we learned? Clearly, there are more

models for understanding professional learning than Guskey's. However, Guskey still matters a great deal. This is evident in the fact that all the other works examined here cite Guskey (2000, 2002, or prior works). In many cases they also explicitly engage with his ideas (more often the 2002 "pathway" model of how teachers learn rather than the more general 2000 model of professional development evaluation). Whether they are attempting to add onto Guskey's framework (King, 2014), re-order it (Desimone, 2009), or question (Clarke & Hollingsworth, 2002; Opfer & Pedder, 2011) or refute the premise of pathway models entirely (Coldwell & Simkins, 2011), subsequent authors are still negotiating with his ideas.

One group of models discussed here aims to capture teacher learning in greater complexity than Guskey. Sometimes that means adding components to what is still essentially a linear model, but more often it means conceptualizing teacher learning as a cycle (e.g., Timperley, 2011) or a complex, nested system (e.g., Opfer & Pedder, 2011) or a process that takes place within particular policy and organizational contexts (e.g., Kennedy, 2014). Those whose response to Guskey was "it's messier than that!" find themselves in good company here.

What specifically do these models add to or complicate about Guskey? Some dig further into the learning activities themselves, identifying features that are associated with effective outcomes (Desimone, 2009) providing taxonomies of different purposes and models (Kennedy, 2014) or drawing attention to the intensity and quality of program "features" (Kennedy, 2016). Guskey's model is somewhat silent on what constitutes a learning experience, and in particular what types of learning experiences are most likely to propel change, so this is a contribution. For educational leaders or other designing or selecting professional learning programs,

these models have practical applications. They also do for those researching or evaluating them.

Another contribution is more emphasis on individual teachers and their motivations for change. Fundamental to many of the models reviewed here is the understanding that change happens in different ways for different individuals (e.g., Clarke & Hollingsworth, 2002; Opfer & Pedder, 2011; Evans, 2014). This obviously has key implications for research and practice (e.g., the design of professional learning activities, or supervision/coaching of teachers as they participate and implement what they learn). Although the teacher is the key actor in the Guskey model, his/her engagement in learning is under-explored (Boylan et al., 2018), and several of these models focus more on antecedents for change, teacher agency, etc. These models raise practical issues for school leaders -- for example, how to give teachers “voice and choice” in professional learning (perhaps while heeding Timperley (2011) that it should not be a free for all) or how to identify and target the outcomes that are most likely to motivate each individual. Looking across the models, another contribution is to put more emphasis on the role of school leaders (e.g., Evans, 2014; Timperley, 2011) and, more broadly, on other forms of organizational or political context (e.g., Kennedy, 2014) within which professional learning takes place. Arguably, these ideas do *exist* at Level 3 within the Guskey (2000) model, but subsequent models elaborate on them more.

I have concentrated here on implications for practice, but these models also have many implications for research or evaluation. Indeed, one consensus across these articles is that professional learning research still needs to be more consistent (e.g., Desimone, 2009), more responsive to context (e.g., Kennedy, 2014), longer-term and

more rigorous (e.g., Kennedy, 2016) and better theorized (e.g., Evans, 2014). The authors emphasize different shortcomings but all seem to agree that research on teacher learning still has a long way to go, and modeling the process is an important part of pushing it forward.

Summary of Key Takeaways

For practice	For research and evaluation
1. Professional learning is complex. Digging deeper into how it occurs can improve design and implementation. More than one model can be informative. 2. Teachers bring different skills, preferences, and motivations to professional learning. Understanding your specific audience is crucial. 3. Professional learning includes affective as well as cognitive and behavioral dimensions. Implementation requires “winning over hearts and minds.” 4. Facilitating professional learning requires skill and capacity. Balancing internal and external resources is key.	5. Professional learning is complex. Developments have occurred, but research still needs more rigor. 6. Such research requires resources: time, funding, technical expertise. 7. Theorizing professional research is worthwhile. It can guide research design and/or provide the basis for empirically testing theories or models. 8. Research, evaluation and policy should move beyond questions of “did it work?” or “what worked” to consider issues of context and subjectivity (e.g., “for whom did it work, where, when, and why? How did teachers understand what they were learning?”

References

- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33 (8), 3 – 15.
- Boylan, M., Coldwell, M., Maxwell, B. & Jordan, J. (2018). Rethinking models of professional learning as tools: A conceptual analysis to inform research and practice. *Professional Development in Education*, 44 (1), 120 – 139.
- Bubb, S. and Earley, P. (2010). *Helping staff develop in schools*. London: SAGE.
- Clarke, D. & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18, 947 - 967.
- Coldwell, M. & Simkins, T. (2011). Level models of continuing professional development evaluation: A grounded review and critique. *Professional Development in Education*, 37 (1), 143 – 157.
- Desimone, L. (2009). Improving impact studies of teachers' professional development: Towards better conceptualizations and measures. *Educational Researcher*, 38 (3), 181 – 199.
- Evans, L. (2014). Leadership for professional development and learning: Enhancing our understanding of how teachers develop. *Cambridge Journal of Education*, 44 (2), 179 – 198.
- Guskey, T.R. (2000). *Evaluating professional development*. Newbury Park, CA: Corwin Press.
- Guskey, T.R. (2002). Professional development and teacher change. *Teachers and Teaching*, 8 (3), 381-391.
- Hall, G.E. and Hord, S.M. (1987). *Change in schools: Facilitating the process*. Albany: State University of New York Press.
- Kennedy, A. (2014). Understanding continuing professional development: The need for theory to impact on policy and practice. *Professional Development in Education*, 40 (5), 688 – 697.
- Kennedy, M.M. (2016.). How does professional development improve teaching? *Review of Educational Research*, 86 (4), 945 – 980.
- King, F. (2014). Evaluating the impact of teacher professional development: An evidence-based framework. *Professional Development in Education*, 40 (1), 89 – 111.

- Kirkpatrick, D. (1959). Techniques for evaluating training programs. *Journal of ASTD*, 13 (11), 1- 13.
- Mouza, C. (2009). Does research-based professional development make a difference? A longitudinal investigation of teacher learning in technology integration. *Teachers College Record*, 111 (5), 1195 – 1241.
- Opfer, V.D. & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, 81 (3) 376 - 407.
- Timperley, H. (2011). *Realizing the power of professional learning*. Maidenhead, England: Open University Press.

Appendix D

ARTIFACT D: PRACTICE BRIEF, IMPROVING PROFESSIONAL LEARNING IN DELAWARE

Introduction

Facing a myriad of pressures from higher academic standards to growing opportunity and achievement gaps, educational systems often turn to professional learning as an improvement strategy. It *can* help. A leading researcher argues that successful reforms almost never occur in its absence (Guskey, 2000). Yet despite the country’s massive investment in professional learning – estimated as high as \$2.5 billion annually or as much as \$18,000 per teacher per year in some districts – there is little evidence of systematic payoff. In addition, teachers typically give the professional learning opportunities available to them mixed if not downright critical reviews (Layton, 2015; TNTP, 2015). This brief draws lessons from Delaware for capitalizing on the promise of high-quality professional learning.

Thanks to decades of research, we know “what works” in professional learning. Certain design features and contextual conditions are associated with teacher and student outcomes (Darling-Hammond, Hyler & Gardner, 2017). Vocabulary has shifted in response. Instead of “professional development” (PD) happening to an educator in a discrete way, we recognize a complex, ongoing process of adult learning with the purpose of promoting student achievement. Adults learn best when tasks are active and highly applicable to their jobs. Instructional change is difficult and takes time: researchers estimate a minimum of 14 hours of activities, and often much more (Yoon et al., 2007). It also requires ongoing support, such as coaching, modeling, and

collaboration. Access to tangible and intangible resources (e.g., time, funding, trust among teachers, support from administrators) can facilitate or constrain learning and application. School and district professional culture matters tremendously. All of these ideas have been encoded in national and state policy, such as the 2015 Every Student Succeeds Act (ESSA) and the Delaware Standards for Professional Learning.

As a state, Delaware has taken several steps towards high-quality professional learning yet there is more work to do. In 2013, the Delaware Teaching, Empowering, Learning and Leading survey revealed mixed perceptions among educators. Most reported that professional learning helped them teach more effectively and improve student learning, yet there were weaknesses in three important areas: differentiation, follow-up on, and evaluation of professional learning (New Teacher Center, 2013). To address these issues and build capacity, the Delaware Department of Education (DDoE) launched the *Reimagining Professional Learning* competitive grant in 2016. Applicants had to address the Standards, meet local needs based on data, and evaluate the impact of their activities on instruction and achievement. Grantees –21 each in 2016 and 2017 – receive ongoing support, technical assistance, and online learning opportunities. Beyond this new initiative, Delaware is also home to several professional learning exemplars.

Vignettes of Promising Practices

Four Delaware programs can help us learn how to design, implement, and lead high-quality professional learning. These programs include the Teacher Leadership Initiative, the Delaware Reading and Writing Project, *Partners4CS*, and the Laurel School District Comprehensive Professional Development Partnership. All four apply the Professional Learning Standards, use research-based designs, and have some

evidence of impact. Together they illustrate the level of intention, resources and follow through that effective professional learning requires. They show how high the bar needs to be, and that it is possible to meet it. The programs highlighted here also vary in important ways. Although small, Delaware demonstrates a variety of innovative approaches to educational problems. These four programs share a common goal of improving instruction and achievement but reach it in different ways. Beyond demonstrating what the standards and research look like in action, each vignette also illustrates a big idea about professional learning. Grappling with these ideas can help us achieve the promise of professional learning in Delaware. The big ideas are: (1) Educators should be agents of their own learning. (2) Balance internal and external expertise. (3) Partnerships enhance learning for everyone, and professional learning can help change educational conditions and (4) Intentional design and leadership support yield powerful results.

1. Educators should be agents of their own learning: Schools That Lead's Teacher Leadership Initiative

As any teacher knows, engagement and motivation inspire learning. Adults in particular learn best when they can direct and make decisions about what they learn and how they will apply it in their own practice (Knowles, 1980). Educators have deep knowledge of their classrooms, their students and the problems of practice they face every day. Professional learning should leverage their knowledge and empower teachers as agents of improvement rather than as passive subjects to be developed. Although this may seem like common sense, it is difficult to “walk the talk” of teacher agency in professional learning. National surveys reveal that district or school administrators make most of the decisions about what, how, and why teachers will

learn (Corwin, Learning Forward and the National Education Association, 2017). In Delaware, we have an exemplar with a different approach: the Teacher Leadership Initiative (TLI), a program of the non-profit Schools That Lead.

TLI brings together cohorts of participating teachers for two years of professional learning, including 14 full-day sessions and ongoing application activities. TLI is structured as an inquiry project, where participating teachers pose questions about their students' learning, then work with a peer(s) to collect data, investigate these questions, and inform their practice. For example, a teacher might want to know what kinds of questions students ask in a literature circle, what type of mathematical discourse students use or how they persist or collaborate in small group work. Participating teachers drive their own development in TLI: they identify what they want to learn and improve about their teaching and collaborate intensively. Over time, participants involve more colleagues in the inquiry cycle, influencing instruction in other classrooms and the school as a whole. As part of TLI, participants can pursue micro-credentials in fifteen different areas of teaching and leading.

Internal and external data suggest positive impacts from TLI. Participants increase confidence in their ability to observe peers and reflect on instruction. They make specific instructional shifts and start to think differently about student learning. In some cases, changes in practice are starting to spread beyond TLI participants to the staff as a whole (Mead and Buttram, 2017). This especially occurs in contexts with a trusting adult culture; school leaders who support the program but do not try to control it; alignment between TLI and other school change efforts; and sufficient resources of time and funding.

2. Balance internal and external expertise: The Delaware Reading and Writing Project

As TLI demonstrates, teachers can provide a wealth of resources for each other. However, ever-changing standards, assessments and technologies mean that some necessary expertise may not exist within schools – yet. Sometimes organizational change requires a jolt from the outside to gain momentum. As they work with each other to change their practice, teachers can benefit from external eyes and additional supports. The Delaware Reading and Writing Project (DRWP) demonstrates how collaboration within school teams can combine with technical expertise and assistance from a university and the state educational agency to promote teacher learning. DRWP is affiliated with the National Writing Project, one of the country’s oldest and best-established professional learning programs in literacy.

The project focuses on resource and curriculum development and spans an entire academic year. Regular and special education teachers from the same school and grade level work together in teams. Teams come together for three in-person and two online sessions led by University of Delaware (UD) and DDoE experts focused on the Common Core standards, text complexity, and the design of performance tasks and formative assessments. Between sessions, they return to their schools and apply these ideas, collaborating to design integrated curriculum modules with support from UD or DDoE liaisons. Teams pilot these modules in their classrooms. At the end of the year, participants present their work and reflect on the experience to the state’s Literacy Coalition and Cadre. The curriculum modules are disseminated statewide.

Evaluation data from the 2016-17 DRWP demonstrate positive results in teaching and learning. Teachers reported knowledge and confidence gains in a variety of content, pedagogy, and assessment topics. Their students also grew as writers.

Writing scores after instruction showed statistically significant improvements in every trait of writing, at every grade level (Mead, 2017).

3. Partnerships enhance learning for everyone, and professional learning can help change educational conditions.: *Partners4CS*

There are many creative ways for universities and public educators to collaborate and build capacity. *Partners4CS* offers an example in a discipline with high demand yet limited background knowledge in most K-12 schools: computer science. To increase students' access to computer science, professional learning for educators is imperative. In 2012, Computer Science and Education faculty at UD received a grant from the National Science Foundation for this purpose. *Partners4CS* offers a paid intensive summer workshop sustained by online sessions during the school year and an annual summit. While leveraging the technical expertise at UD, the project is designed creatively so that learning happens continually and for everyone. For example, it includes a course through which UD computer science undergraduates offer ongoing support and resources to participating teachers and their students, and in the process gain valuable field experience. Project investigators are publicly engaged in their own development and research about professional learning, and regularly publish and report on the project.

So far *Partners4CS* has reached 96 teachers in seven districts and 25 schools, who together teach approximately 1000 students. The goal is for participating teachers to implement either a full computer science course or integrate computational thinking into other STEM courses. Teachers report learning about computer science and strategies for teaching it, and most transfer this new knowledge into practice (Pollock et al, 2017). Middle school students taught by *Partners4CS* participants also

demonstrate positive changes in their knowledge of, and attitude towards, computing (Mouza, Marzocchi, Pan & Pollock, 2017). Partners4CS has also contributed to structural changes, such as the establishment of a state chapter of the Computer Science Teachers Association and a Career and Technical Education pathway for computer science. In 2017, Governor Carney signed House Bill 15, requiring that all public high schools offer at least one class in the subject. This law represents progress increasing access and raising awareness about computer science in Delaware and demonstrates the potential role of professional learning in policy.

4. Intentional design and leadership support yield powerful results: Laurel School District Comprehensive Professional Development Partnership

All three programs profiled so far involve teachers from around the state, often working in school-based teams. Sometimes, an individual organization has specific needs for professional learning identified either internally or externally, such as through an accountability system. In 2015, Laurel Middle School in a rural county in southern Delaware was identified as a “Focus School” due to persistently low performance on state test scores. District leaders collaborated with professionals at UD to respond to the school’s student and teacher learning needs. Since that time, the initiative has expanded to all four schools in Laurel. The partnership demonstrates the power of intentional design, coherence, and leadership engagement in professional learning.

Two centers at UD – the Professional Development Center for Educators (PDCE) and the Delaware Academy for School Leadership (DASL) – provide coaches in all content areas including special education, and at all grade levels. Professional learning days align with district instructional priorities and include at least five days

per year. District leaders from superintendent down strongly invest in this partnership, present at and participate in these days. School leaders have sessions specific to their needs, so they understand new instructional practices and can support the teachers in their buildings. Between sessions, coaches visit schools at least monthly for targeted and responsive services. For example, they may work individually with teachers, facilitate PLCs, support instructional planning, or conduct walkthroughs. While maintaining confidentiality, coaches report monthly about trends observed. Together, educational leaders and coaches identify priorities and adjust plans based on data.

This strong infusion of research-based professional learning has yielded results in Laurel, especially in the middle grades where the partnership is most established. Between 2015 and 2017, Laurel Middle School improved achievement at all grade levels, increasing its math state test scores by 24 percentage points and English Language Arts by 19 points schoolwide. The school's turnaround and professional learning journey were featured in a public television story (Delaware Department of Education, 2017; Barrish, 2018).

Recommendations

These vignettes illustrate four different possibilities for what high-quality professional learning can look like. They are not the only examples in our state but they are also not yet typical. How can we multiply research-based professional learning opportunities, so all Delaware teachers and students can benefit? What challenges still exist?

Looking across the vignettes, we see two complex challenges. The first is finding resources for intensive, sustained professional learning. In different ways, all four programs require an investment of both money and time. This includes not only

time for a teacher to leave the classroom to “attend PD” but also to work to implement new ideas in the classroom, and for her colleagues and leaders to engage in and support that process. We must become more informed and realistic about what is required for changing practice, and thus prioritize. Any individual teacher cannot pursue every opportunity; attending both TLI and DWRP might be too much. Likewise, choices must be made on the school or district level. Educational leaders play an important role in prioritizing professional learning, and sometimes buffering external demands.

This leads to a second challenge: bringing the big ideas about professional learning together. A skeptical reader might notice tensions between the vignettes. For instance, TLI and Laurel demonstrate different beliefs about teacher choice in professional learning. Where TLI equips teachers to choose their own inquiry, Laurel activities are coordinated. It is complex but not impossible to combine teacher agency with organizational alignment. It requires helping teachers recognize school/district needs and encouraging collaborative responses to those needs. It also requires more careful attention to the process of adult learning. Even centralized initiatives can create opportunities for individual teachers to make choices and express needs. Ultimately, leading high-quality professional learning requires complex and strategic balances.

To make exemplary professional learning like that illustrated in these vignettes more widely available, we offer the following recommendations.

Teachers and teacher leaders should exert agency in their own professional learning, speaking up about what they need and value in it. Using high quality examples such as those discussed here, they should pursue opportunities that will be

worth their time and supportive of their growth. They should also contribute internal expertise to colleagues in job-embedded learning activities.

School leaders should act as intermediaries between the professional learning needs of their teachers and the resources of their districts. This may require focusing and buffering. As Laurel demonstrates, leaders should be as involved as possible in learning alongside their teachers, then support ongoing implementation. They should invite teachers to voice professional learning interests. Finally, leaders should understand their influence over the context in which professional learning takes place.

District leaders, state leaders and policymakers should reflect on and apply the lessons from these vignettes and the standards, assessing honestly whether current professional learning practices and conditions align. If not, they should take action to close gaps. They should advocate for resources, honor in-house expertise and bring in necessary external perspectives, and work to increase understanding of high-quality professional learning.

University partners should learn from the examples in this brief about how partnerships can build capacity for professional learning. They can ensure that programs they design or propose align with the research on professional learning. Then they can add to that research base by sharing insights and results from their work.

References

Barish, C. (2018, February 12). Low performing school rebounds with academic rigor, investment in teachers. *WHYY*. Retrieved February 14, 2018 from: <https://whyy.org/articles/low-performing-school-rebounds-with-academic-rigor-investment-in-teachers/>

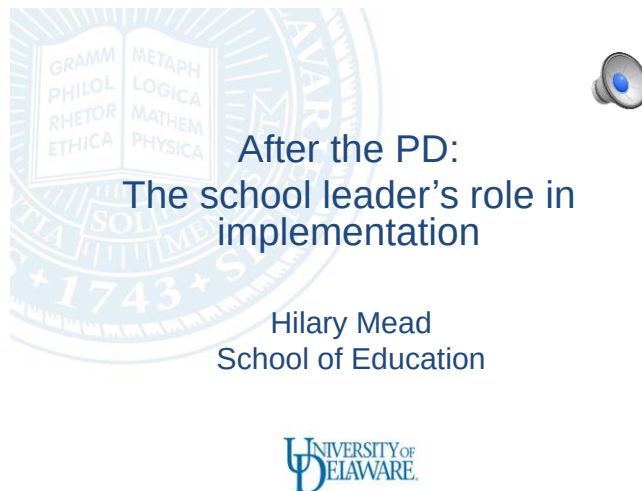
- Corwin, Learning Forward & National Education Association (2017). *The state of teacher professional learning: Results of a nationwide survey*. Retrieved October 21, 2017 from https://dashboard.learningforward.org/wp-content/uploads/sites/12/2017/08/the-state-of-teacher-professional-learning-2017.pdf?_ga=2.252399947.373935782.1508519059-2068151134.1502121936
- Darling-Hammond, L., Hyler, M.E. & Gardner M. *Effective Teacher Professional Development*. Palo Alto, CA: Learning Policy Institute. Retrieved October 29, 2017 from: https://learningpolicyinstitute.org/sites/default/files/product-files/Effective_Teacher_Professional_Development_REPORT.pdf
- Delaware Department of Education (2017). *State assessment results show more students proficient in math and English*. (Press release). Retrieved October 29, 2017 from <https://www.doe.k12.de.us/site/default.aspx?PageType=3&DomainID=4&ModuleInstanceID=20&ViewID=6446EE88-D30C-497E-9316-3F8874B3E108&RenderLoc=0&FlexDataID=20716&PageID=1>
- Guskey, T. R. (2000). *Evaluating professional development*. Newbury Park, CA: Corwin Press.
- Knowles, M.S. (1908). *The modern practice of adult education: From pedagogy to andragogy*. Wilton, CT: Association Press.
- Layton, L. (2015, August 4). Study: Billions of dollars in annual teacher training is largely a waste. The Washington Post. Retrieved November 17, 2016 from: https://www.washingtonpost.com/local/education/study-billionsof-dollars-in-annual-teacher-training-is-largely-awaste/2015/08/03/c4e1f322-39ff-11e5-9c2ded991d848c48_story.html
- Mead, H. (2017). Delaware Reading and Writing Project 2016 – 17 Data analysis. University of Delaware.
- Mead, H. & Buttram, J. (2017). Study of the Teacher Leadership Initiative. Newark, DE: Delaware Education Research & Development Center.
- Mouza, C., Marzocchi, A., Pan, Y., & Pollock, L. (2016). Development, implementation and outcomes of an equitable computer science after-school program: Findings from middle-school students. *Journal of Research on Technology in Education*, 48 (2), 84 – 104.
- New Teacher Center (2013). TELL Delaware results. Retrieved October 23, 2016 from: <https://telledelaware.org/results/report/80/1346#PD>

- Pollock, L., Mouza, C., Czika, A., Little, A., Coffey, D. & Buttram, J. (2017). *From Professional Development to the Classroom: Findings from CS K-12 Teachers*. In Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education (SIGCSE '17). ACM, New York, NY, USA, 477-482.
- TNTP (2015). The mirage: Confronting the hard truth about our quest for teacher development. Retrieved November 22, 2016 from:
<http://tntp.org/publications/view/the-mirage-confronting-the-truth-about-our-quest-for-teacher-development>
- Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>

Appendix E

ARTIFACT E: AFTER THE PD: THE ROLE OF SCHOOL LEADERS IN IMPLEMENTATION (NARRATED POWERPOINT)

Slide 1



This is Hilary Mead from the University of Delaware. I am a doctoral student in the educational leadership program here and I also work for two centers at UD. I've been involved in school leadership and professional development programs for the Delaware Academy for School Leadership (DASL) and I have conducted program evaluations and research for the Delaware Education R&D Center. I am fascinated by the connections between all these areas. How can we use lessons from research to improve professional learning, and what is the role of school leaders in that process?

Slide 2



Purpose



- Share key ideas from research
- Generate implications for practice
- Improve implementation of PD
 - Better return on investment

In this presentation I am going to synthesize some of what I have learned and, I hope, make it useful to you. My overall goal is to give you insights and tools to improve the implementation of professional learning in your schools and settings, in other words how much it actually moves teacher instruction and student learning – the return on investment, so to speak. We’re going to focus on what you or other school leaders (principals, APs, teacher leaders) can do in that implementation process.

Slide 3



Ah, PD days. Probably you just had some at the beginning of the year, or are planning some in the future. Here are six different perspectives on those days. Notice the bottom right picture. Now, this may not be true in your school or district. But mounds of research and lived experience tell us that most PD is boring and not always effective. We treat teachers like we would never treat students.

Let me add a quick note about vocabulary. For ages, we have talked about professional development or PD for short. More recently, it has been called professional learning. That's a better phrase to capture the ideas (all proven in research) that teacher learning is ongoing and requires active engagement. It happens not just in workshops, classes, and in-service sessions but every day in formal as well as informal interactions. In this presentation I'm going to try to remember to use "professional learning" but I might sometimes fall back on old habits and use PD.

So, this meme shows us what six groups think about professional learning. But there is one group missing. [click] We're going to focus on the role and perceptions

and actions of school leaders here. [click] As a school leader, do you attend these sessions? Where are you, physically? What about mentally? What do you do DURING professional learning? And what about after it?

Slide 4



Big idea #1: It is part of your job

Professional Standards for Educational Leaders /
Delaware Administrator Standards 6 & 7

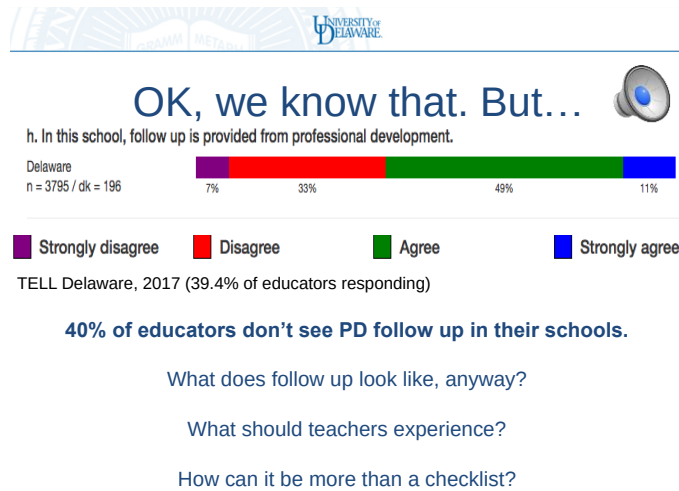
Effective educational leaders...

- *Develop the professional capacity and practice of school personnel*
- *Foster a professional community of teachers and other professional staff*

...to promote each student's academic success and well being.

Let's start with a big, basic idea. Leading professional learning is a big part of a school leader's job. Click to see what two Delaware School Leader Standards have to say on that subject. Many of the other standards pertain to professional learning as well.

Slide 5



I know, that wasn't news to you. I do a lot to help my teachers learn, you're thinking. But here's a tricky reality: staff members in your school don't necessarily agree. Or they might recognize that professional learning activities happen, but think that those activities come and go without much follow through. Click to bring up some data. Last spring, almost 40% of the educator workforce in Delaware took a survey. One big section was about what they called PD. Here's what they said about the statement: "in this school, follow up is provided from professional development."

Red and purple are disagree and strongly disagree. 40% of educators don't see follow up. Only 11% gave the most positive rating. This is an area in which we can do better. Tons of national research backs up the idea that implementation is the weak link in the professional learning. We plan good activities, we facilitate them well or get others to do it...and then those efforts don't bear fruit. I believe that educational leaders know they are supposed to do something "after the PD". I can't tell you how many times at UD I wrote reports that included statements like, "school leaders must monitor the translation of PD into practice." But what does this look like? What should

teachers see and feel? And how can you do it in a way that goes beyond walking around with a checklist? That might get you compliance, but it won't get you deep and personalized teacher learning.

Slide 6



Big idea #2: Leadership matters

- Differences in implementation of PD initiatives are often associated with school leadership.
- Do not underestimate how many resources implementation will require.
- Your practice may need to shift, too.

Here's another big idea from research I have read, and conducted. It's pretty simple. You matter. School leaders matter. A lot of times at UD we evaluate the implementation of an initiative in many settings. Maybe a district is trying out a new instructional approach in multiple schools. Usually, some schools experience better results than others. When we look deeply into why that is, many times school leadership is part of the answer. Along with this power comes a burden. Implementing professional learning effectively takes a lot of time – and most of it has to be invested *after* the activity itself. Don't underestimate the resources required on the back end. Also recognize that you yourself – what you do every day, how you view and think

about instruction, student learning, even your own job – those things might also need to shift so you can support professional learning.

Slide 7



Review of research

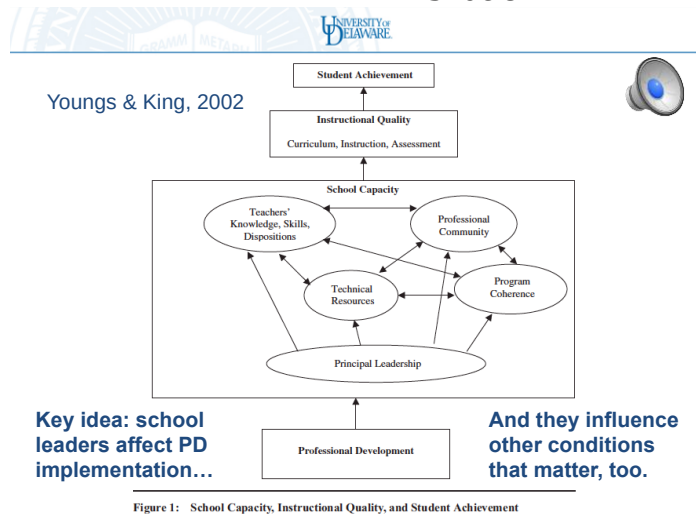


- Empirical
- Peer reviewed
- Theories/frameworks of teacher learning and implications for school leaders
- Case studies of leader's role in implementation of professional learning

Still on board? That's awesome. In my roles as a UD student and researcher of professional learning, I've read and written lots and lots of studies. And I know you have many more urgent demands on your time, so I'm going to give you the highlights of what I learned, to crunch a body of research and give you some takeaways for what it means for your work every day. Just so you know, there is a list of the studies that went into this work at the end of this presentation. These are empirical studies that mostly came from peer-reviewed publications, usually academic journals. I am including some studies that I and my colleagues and I conducted of Delaware programs, because I think the local lens is useful. I'm always interested in, and I think school leaders can always use, illustrations and examples of what it looks like to practice leadership for professional learning implementation. So many of the studies I'll be talking about are case studies that go in depth into a particular story or situation.

I also include some quantitative work as well as some theoretical pieces that try to build a model of what happens when teachers learn.

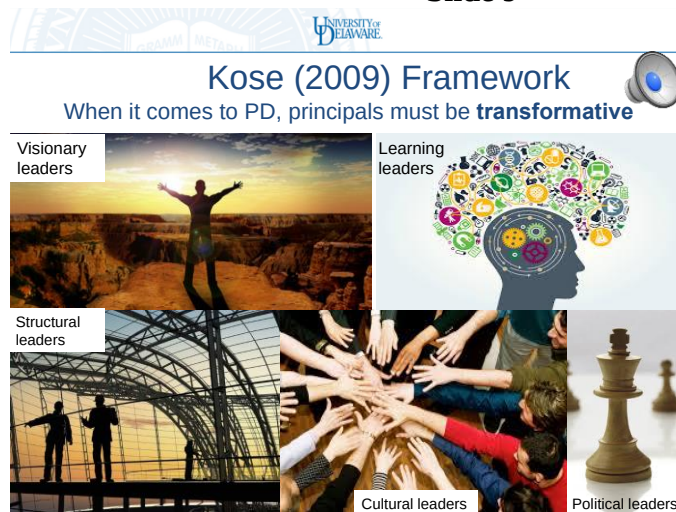
Slide 8



Here's an example of what that kind of model might look like. This comes from an article by Youngs and King in 2002, and it's pretty complicated. Usually models like this they have lots of arrows and boxes like this. This one should be read from bottom to top. It basically shows what happens when professional development comes into a school to affect that big box, school capacity, and then from that comes changes to instructional quality and at the very top, the ultimate goal, is changes to student achievement. There are two key things I want you to take away from this model. **[Click]**. The first is that principal leadership matters. This shows graphically what we just said. Principal leadership is that big circle in the bottom of the middle box. What principals do has an influence on school capacity and on transforming PD inputs into improvements in instruction and student learning. And the other thing is that principals also influence other influences. If you look at all of those other circles,

which are conditions in the organization that we'll get more into, principals have a lot of leverage on things like the professional community or coherence in a school. So principals matter themselves and they matter because they control other things that matter.

Slide 9



Now I want to look at a second framework, which is different in its set up. This comes from a study by Brad Kose in 2009 where he looked specifically at the actions school leaders took to help their teachers become more socially just and to provide more equitable instruction. He identified five different roles for professional learning leaders to assist with implementation. I've illustrated each with an image even though he doesn't have a visual. And you can click through them as I go. Kose says that principals must be transformative visionary leaders, they need to have a vision for the professional learning. They need to be learning leaders, they themselves are actively involved in professional learning. They need to be structural leaders. In order to make professional learning stick and be implemented, there are many structural concerns. Cultural leaders. They are also many cultural concerns, and the culture in the building

can either get in the way of changes to professional practice or really facilitate. And finally, political leaders. In order to execute change in the school of the kind that professional learning may require, there are many political roles that school leaders must play.

Slide 10



Brainstorm



Think about a current or planned professional learning experience or initiative in your setting.

Write a one-sentence summary of its purpose.

What are you currently doing to follow up with teachers and support implementation?

In a minute, we'll go through each of the five roles suggested by Kose and talk in more depth about what that looks like. And I want to just acknowledge that I am using Kose's framework in broader way than he intended in his study. I'm doing that because I think those five roles are a great organizer for thinking about leadership for implementation. And I think that he would be ok with that. Before we get there, I want to take a minute to harvest some thoughts about your own setting. Take a piece of paper or a document and think about a professional learning activity or initiative in your setting that is happening now or that you're planning. It doesn't have to be schoolwide. It might only involve a small group of people, but it should be something that you hope will have ripple effects on the organization. I want you to just in one sentence summarize the purpose of that professional learning. Then I want you to

brainstorm all the things you are currently doing to follow up on that activity and support implementation. After you get that list, as we go forward in the five roles for professional learning, please think about whether any of the actions you listed could be categorized into those roles. And I want you to think about what additional practices you could add to your list so you might be able to do more to support the implementation of professional learning.

Slide 11



Visionary leadership for professional learning



1. Communicate vision for what professional learning can accomplish in the school.
2. Align professional learning with school vision/goals.

What does it look like?

What could you do?

What's involved with being a visionary leader of professional learning?

Research suggests two big ideas, and I invite you to click through them with me. You need to communicate a compelling vision of *why* teachers should implement the new practices in the first place. Why are you doing the professional learning? And second you need to align and connect the different initiatives that are happening in the school to each other and to that vision. Kose gives an example of the first idea. He studied the principal of a rapidly diversifying school in the Midwest. In this school, a vision of equity was not just lip service. The principal made it very specific and practical and connected to the professional learning that was provided. For example, she said that

there would be an ongoing learning activity about Hmong students, and she framed that as an expression of the belief that students from all backgrounds can learn at high levels. Here's a Delaware example of the second idea. A principal in this state regularly (a couple of times a year) communicates to her faculty about how all the professional learning activities in the school are connected to each other and to the school vision. She does this in words and visuals. She also shares data regularly to show the progress that is being made.

Those are some ideas. What could YOU do to connect professional learning to your school vision, and to communicate that to stakeholders? Jot down some ideas at this point.

Slide 12



Learning leadership for professional learning

1. The more *you* learn, the more you can support implementation.
2. Differentiate for teachers (as you would for students).
3. Optimize internal and external expertise for learning.

What does it look like?

What could you do?

Now we look at learning leadership for professional learning. I know it sounds circular. There are three ideas in this role I'd like to emphasize. And the first is so much easier for me to say than for you to find time to do, and I recognize that. But it is a simple truth that the more school leaders know, the better able they are to support implementation of professional learning. You've heard of content knowledge and probably also pedagogical content knowledge, which is what teachers need to know to teach effectively. There's also a construct called leadership content knowledge, what school leaders need to know to be able to successfully lead and guide a program. You need to know *enough* about the what and the how of the professional learning: writing or STEM or classroom management, whatever it is.

How are you supposed to get that information? Well, we hear from teachers that they do notice when school leaders participate in professional learning. By showing up, you send a very clear signal about your priorities, and you establish credibility for all your future follow-up efforts. Of course, you can't go to everything. But one test of whether something really can be a professional learning priority in the school is to ask: can you make some time for it? Other strategies include empowering others, instructional leaders, teacher leaders, to attend sessions and bring back key points to you. And you can also demonstrate curiosity and talk to teachers about what they are learning and try to learn from them.

Second, another big idea is to think about yourself as the teacher of a group of adults. You are responsible for making sure they learn what they are supposed to and use what they have learned. But just like a teacher wouldn't use the same lesson plan for every student, you also need to differentiate, not just in how teachers experience professional learning, but in what happens afterwards. Some will require a lot more

resources (more coaching, more visits, more modeling, more walkthroughs) than others.

Last, research is clear that when you want to support teacher learning, it takes a mix of internal and external resources. So absolutely, use teacher leaders and other experts in your building as much as you can to plan, facilitate and follow-up from professional learning. That definitely helps with teacher buy in. But you also have to be realistic about t capacity. Do you have people in your building who already know how to use new practices well, and to teach others how to do it? And do they have the time to do so? Research suggests that sometimes an external voice or set of eyes or hands is needed to really spur change.

How can you take these ideas about learning leadership to heart? What can you try?

Slide 13



Structural leadership for professional learning



1. Provide resources for implementation: **time**, coverage, materials/technology, funding, space, etc.
2. Make structural changes to support new practices.

What does it look like?

What could you do?

If you ask any teacher what prevents him or her from using what they learned in professional learning, you're going to hear about structures. Probably #1 you'll hear

about time or the lack of it. Time to go to PD, time to visit other teachers, time to collaborate and meet and look at student data. There are other issues associated with that. I need someone to cover my class, I need different materials, we need better technology, we need more budget, we need a space to meet. All different kinds of issues. As a structural leader, the work is to find those resources, getting creative if necessary, and to change the structures in the school to support new practices. Here are several examples from the research about what that looks like. School leaders have reworked the master schedule to promote peer observations. They have changed the way students are assigned to classes to teachers share a group of students. That way they can more meaningfully engage in inquiry and PD about student learning. Some have created new teacher leadership roles or structures to help support implementation of new practices. If you need someone to spend time visiting and coaching and following up in the classroom, maybe that person can be an existing teacher. Simply put, school leaders can cover classes in a planned fashion or even spur of the moment. I had a teacher say to me that she knew that her school leader, you could call her up and she'd be down the hall to cover the class so that the teacher could get out and see something.

Lastly, get creative about finding time and sometimes that might mean giving teachers permission not to do other things, to take things off the plate. I have an example of a principal who decided it would be acceptable for a team of teachers to step out early from the winter concert so they could have precious time to meet and do some lesson study. So think about what structures you might need to change or to create so your professional learning can take root and flourish in your school. What specifically can you do?

Slide 14



Cultural leadership for professional learning



1. Build trust and cultivate collaboration.
2. Balance support and accountability.
3. Model what you seek in professional learning.

What does it look like?

What could you do?

Now we are going to talk about cultural leadership and there are three points I want to emphasize. The first is very very strong in research and makes a lot of sense to anyone who has spent a lot of time in schools. And that is that trust and collaboration are huge. So if you want to ask a teacher to change, to discard old practices, to try some new ones, to take risks, perhaps to fail, that person must have trust in the school leader and in her colleagues. And those things relate very strongly to each other. So school leaders influence greatly the way that teachers relate to each other as peers and professionals in the building. Some of that is structure. Principals have much influence over the time that teachers have to work together and build relationships, but some of it is about the modeling and the messages that are sent from leadership about the importance of collaboration.

Second: you want to balance support and accountability in following up from professional learning. So yes, you want to set the expectation that if resources – time, money and coverage, have been invested in someone going to professional learning – that they should bring something back to benefit the school. And you might want to do walkthroughs or to be very specific with your staff about new practices that you are

expecting to see in the classroom. But with that needs to come that deeper support that we talked about, of understanding that it might not look pretty at first, or that it might require additional resources, or understanding that they are taking a personal risk and putting themselves out there. One way to show support is to highlight the work that teachers are doing when engaged in professional learning. Especially if it's not a full-school initiative but it's a small group who are trying something, you can affirm that effort and create peer support by giving them opportunities to share out to the whole staff whether in newsletters, meetings or other channels, and that will help create the momentum and support to change practice.

The last point is again easier to say than to do but it's to remember that teachers are watching you. School leaders need to model whatever they want to have happen in professional learning. So if the activity is all about growth mindset or critical thinking, you should be trying to demonstrate those behaviors too. Kose has an example of this. Remember, it was professional learning all focused on equity within the school. And a teacher was talking about her principal and how she was so willing to ask tough questions about student equity and to really engage in learning. She used the phrase: "my principal pushes herself to wonder." So, what would your teachers say about what you are modeling and how well it aligns with what you are emphasizing in the professional learning?

Slide 15



Political leadership for professional learning



1. Prioritize. Buffer teachers from distractions.
2. Advocate for professional learning resources.
3. "Tap" participants strategically.

What does it look like?

What could you do?

Last but not least we come to political leadership for professional learning. There are three ideas here I would like to illustrate. The first is to prioritize and buffer. So if you are going to be devoting as many resources to professional learning as you need to, there's only going to be so many initiatives that your building can handle. So part of successful implementation is knowing when to say no, whether that is saying no to an individual teacher who wants to do something that isn't aligned, or to a district office that would like to get you involved in one other initiative, or to a previous practice that no longer fits or is needed. So an example of that is a principal who stopped using a particular walkthrough practice because he or she felt that it was inconsistent with the new direction that the professional learning was leading them in.

Second, there's the role of advocacy for more resources. Whether that's negotiating with your district for more building-level professional time or writing grants for more financial resources, or getting creative in a whole variety of ways, school leaders can play a role in expanding the pie of what's available for professional learning. Third, also be thinking about politics on the staff level. So a lot of times a

place where this comes in is in thinking about who to encourage to do what professional learning activity. Many times even if an activity is voluntary for teachers, a school leader plays a role in tapping or guiding certain teachers to that activity. And we've learned that the match there is really important. Especially if you are trying to bring something new or to change practice in a significant way, that thinking about the politics of who your early adopters need to be is an important consideration. So now I want you to think and look at your list of practices and brainstorm: whose support will you need for implementation, and how can you obtain it?

Slide 16



Questions? Feedback? 

Hilary Mead
hmead@udel.edu
978-760-9125

So we have made it through these five roles that school leaders play in order to implement and follow up on and support professional learning in their schools. Again, they were visionary, learning, structural, cultural and political. And for each one we looked at more specific actions or ideas and I tried to illustrate with concrete examples. The two big ideas I want to leave you with are that this work of “after the PD,” what happens when the teachers come home, when the workshop is over, when the coaching has concluded, how you can make sure things stick, this is one of the

most important things you do. It's hard. It takes a lot of resources. None of the suggestions I have made are easy and I recognize that. But you have a lot of power and influence over making sure the investments, human and financial, that are made in professional learning will pay off in your school. Here's my contact information. I welcome your feedback about any of the ideas or further conversation about what it might look like. The next slide will list the studies that are synthesized in this presentation. Thank you so much for listening.

Slide 17



Studies Consulted

Bredeson, P.V. & Johansson, O. (2000). The school principal's role in teacher professional development. *Journal of In-Service Education*, 26 (2), 385 – 401.

Dierking, R.C. & Fox, R.F. (2012). "Changing the way I teach": Building teacher knowledge, confidence, and autonomy. *Journal of Teacher Education*, 84 (2), 129 – 144.

Evans, L. (2014). Leadership for professional development and learning: Enhancing our understanding of how teachers develop. *Cambridge Journal of Education*, 44 (2), 179 – 198.

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

King, F. (2016). Teacher professional development to support teacher professional learning: Systemic factors from Irish case studies. *Teacher Development*, 20 (4), 574 – 594.

Kose, B.W. (2009). The principal's role in professional development for social justice: An empirically-based transformative framework. *Urban Education*, 44 (6), 628 – 663.

Olsen, H.S. (2010). How leadership content knowledge in writing influences leadership practice in elementary schools. (Unpublished doctoral dissertation). University of California, Berkeley.

Supovitz, J., Sirinides, P. & May, H. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly*, 46 (1), 31 – 56.

Stosich, E.L. (2016). Building teacher and school capacity to teach to ambitious standards in high poverty schools. *Teaching and Teacher Education*, 58, 43 – 53.

Youngs, P. & King, M.B. (2002). Principal leadership for professional development to build school capacity. *Educational Administration Quarterly*, 38 (5), 643 – 670.

Appendix F

ARTIFACT F: *DELAWARE DEPARTMENT OF EDUCATION SPECIFIC AND INNOVATIVE IMPROVEMENT GRANT META-EVALUATION*

Delaware Department of Education Specific and Innovative Improvement Practices Grant (SIIP) Metaevaluation

Joan L. Buttram, Ph.D.

Hilary Mead, M.P.A.

Sara Shaw, M.S.

Delaware Education Research & Development Center
University of Delaware
Newark, DE 19716



September, 2014

PUBLICATION T2014.24

Introduction

Fourteen projects were funded through the Delaware Department of Education (DE DoE) Specific and Innovative Improvement Practices Grant (SIIP) for the 2013-14 school year.¹ These innovative and promising projects focused on one or more of the following goal areas:

- teacher-led projects that drive improved student outcomes,
- Common Core implementation and assessment,
- student supports and dramatically improved school climate², and/or
- accelerating the achievement of under-performing groups.

The Delaware Education Research & Development Center (DERDC) at the University of Delaware conducted an evaluation of the SIIP projects. We developed 16 evaluation rubrics³ to assess the implementation and outcomes of the individual projects as well as the attainment of one or more SIIP overarching goals (see Appendix A); met with individual projects to review the rubrics and evaluation procedures; provided technical assistance to projects when requested or appropriate; and assessed the 14 SIIP projects using individual project-generated final reports and the evaluation rubrics. In essence, we were conducting metaevaluations, using the 14 projects reports as our data source. We also consulted with DE DoE officials throughout the year.

Evaluation Findings

The evaluation findings are presented from two perspectives. We first summarize the evaluation findings by project and then look across the 14 projects to identify broader trends.

Evaluation Findings by Project

Brief summaries below highlight the evaluation findings for each project; the SIIP projects are listed in descending order based on their evaluation rubric total score. Each rubric was scored on a five-point scale with a “3” anchoring “expected performance.” These summaries are followed by Table 1 that presents the individual projects’ scores on the evaluation rubrics.

¹ Brief descriptions of the individual projects can be found at <http://www.doe.k12.de.us/senate148/siip.shtm>.

² This goal was divided into two sub-parts: the first focused on improving climate for students and the second for improving climate for teachers.

³ The 16 evaluation rubrics were divided into three sets: process (7 rubrics), outcome (5 rubrics), and then goal-specific (4 rubrics). Each project’s final report received scores for all of the process and outcome rubrics as well as rubrics for the specific-goals that they addressed. If projects addressed more than one goal, a score was entered for each relevant rubric and then averaged to make the 13th score.

SIIP Grantee/Project	Evaluation Summary
BRINC (Brandywine, Indian River, New Castle Co. Vo-Tech, and Colonial), Linking to the Future Evaluation Total Score=40	This project formed a four-district consortium, contracted with a Learning Management System and a blended learning expert, and provided professional development (PD) for 40 teachers in 10 high schools. This Year 1 grant was part of a more extensive vision and plan for BRINC. The project had extremely strong district support, and dissemination and future planning were notable strengths. Information about the engagement of educator and student participants is provided. Application of PD was uneven; not all 40 teachers taught blended learning units with fidelity. Little information is included about student achievement outcomes.
All Red Clay Elementary and Middle Schools (Red Clay), 4 th to 6 th Grade Partnership Institute Evaluation Total Score=37	This project engaged all 167 4 th , 5 th , and 6 th grade teachers in a yearlong PD experience focused on fractions and proportional relationships. This ambitious plan was executed as intended and some positive student achievement results were seen; students scored relatively higher on questions related to fractions than on the DCAS as a whole. The project was a consistent performer, scoring near the expected level (at least a 2) on all 13 rubrics. The district devoted substantial resources to this initiative. Dissemination occurred both statewide and nationally. The largest challenge for this project was monitoring and supporting teachers' use of new knowledge and skills. Although the project lead attempted to do walkthroughs, the task exceeded the capacity.
Stanton Middle School (Red Clay), Implementing a Trauma-Informed System of Care for Stanton Middle School Evaluation Total Score=35	This project included the use of a universal screener, social-emotional and behavioral interventions for students, PD and coaching for teachers by the Student Support Team. Organizers monitored the project using a variety of data sources. Acquisition of resources and dissemination were also strengths. Some aspects of the project did not occur as planned (e.g., students were screened two times rather than four). Student outcome data were mixed and evidence about changes to educator practice was limited.
Brandywine, Concord, and Mount Pleasant HS (Brandywine), Increasing Student College and Career Readiness in the Courses of Biology and AP Biology Evaluation Total Score=34.5	This project adopted the Science and Global Issues curriculum; funds were used to purchase aligned textbooks. Teachers did not acquire supplemental teacher kits until mid-year, slowing implementation. Teachers received PD and worked in PLCs to co-plan lessons; administrators also participated in a district-wide PLC to support their supervision of science. Student enrollment from under-represented groups in honors and AP Biology courses increased. Meanwhile, the number of Biology course failures (overall, and for both African American and special education subgroups) fell. Science DCAS scores grew slightly but did not reach the target. Limited dissemination of practices or results occurred.
Gallaher Elementary School (Christina), GOALS (Gallaher's Outstanding Approach to Learning and Student Improvement) Program	Gallaher fine-tuned its RTI processes by hiring additional interventionists, providing teachers with literacy toolkits, subscribing to web-based intervention programs, and embedding more data review into PLCs. The project was able to extend its scope by obtaining more resources than intended and providing a model to the district's RTI committee. However, no dissemination outside of Christina was described. There were some delays posting positions and hiring interventionists. Although the report

Evaluation Total Score=32	states that student growth occurred and some targets were met, the interpretation of these data is not clear. Similarly, there is limited information about how classroom teachers changed their practices in Tier I or II instruction although improvements are claimed.
POLYTECH High School (Polytech), SIIP at POLYTECH HS Evaluation Total Score=31.5	This project comprised a wide range of activities, broadly linked to improving both academic and career/technical education (CTE) to meet new standards. These included creating an enrichment period, providing PD on the CCSS, acquiring technology and CCSS-aligned resources, visiting other CTE programs, and increasing the number of CTE certifications. Overall the project proceeded as intended although there was limited information about how monitoring the <i>grant</i> was distinct from school oversight in general, and what data were included. Student and teacher outcome results were weaker, partly because the measures were not targeted. This project faced a challenge due to leadership transitions and school culture; participating educators were not ready for instructional follow up, walkthroughs, etc. Dissemination was a strength.
H. O. Brittingham and Shields Elementary Schools (Cape Henlopen), Bridging the Gaps -through Blended Learning Evaluation Total Score=30.5	This project provided PD to fifth grade math teachers in two diverse schools in order to promote collaboration through blended learning practices. Due to snow days, one of the two schools fell behind in the pacing, which limited collaboration between the classrooms. Additionally, technology issues at H. O. B. hindered the blended learning process. While evidence demonstrating that PD occurred, there is no evidence supporting how the PD was implemented in the classrooms. However, the project provided evidence dissemination at state and national conferences showcasing their work.
North Laurel Elementary and Laurel Intermediate Schools (Laurel), i-Impact Evaluation Total Score=24.5	This project purchased i-Pad minis to provide tiered instruction, intervention, and enrichment to meet student needs. A communication breakdown in the district office delayed their purchase which, in turn, meant that teachers and students used the new technology for only 1.5 marking periods. Although all grades 2-6 teachers received training on using the iPad minis, no empirical evidence is presented on the quality of the training, teacher use of the iPad minis during instruction, or increased student engagement. Inconsistent internet access and insufficient numbers of iPad minis also hindered the project. Student achievement targets were not met. Plans for continued use of the iPad minis next year and dissemination were incomplete.
A.I. du Pont, Dickinson, McKean, Conrad, Cab Calloway, Delaware Military Academy, and Wilmington Charter HS (Red Clay), 11 th Grade SAT Prep Evening Program Evaluation Total Score=23.5	This project provided an SAT prep program to Red Clay School District students and funded improvements to the program (e.g., offering sessions at more sites/times and offering transportation to students). It is not clear from the report what PD was delivered. Additionally, the project provided limited evidence of monitoring outside of the project costs. There was no evidence of dissemination of project successes outside of those involved with the project. However, the project was successful at reaching its goals outlined for student achievement, as participants increased their SAT scores above the PSAT predictions.

Warner Elementary School (Red Clay), WIZ (Warner Inspiration Zone) Evaluation Total Score=22.5	This project provided opportunities for students to engage in a variety of clubs in an elementary school. Additionally, this project provided opportunities to teachers to promote leadership. However, the project was not successful in expanding the WIZ model to engage second graders, as originally planned. Therefore the project did not recruit sufficient students and educators. There was also limited evidence of district support on this project. There was no evidence of dissemination outside of the project's school and future plans were incomplete.
Chipman Middle School (Lake Forest), Engaging Students in History through Common Core Evaluation Total Score=22	This project took place in an eighth grade social studies classroom. Grant funds purchased a mobile computer lab and materials to increase access to primary historical sources. The project extended beyond original plans and involved an additional group of students in a "skill review" class. While students found the activities engaging, they struggled with some of the skills required, and the snow interruptions did not help. Student assessment results did not reach the growth targets and were not aligned to state requirements. District support, monitoring, and dissemination were lacking, and the teacher has been re-assigned to a different grade level. These results raise questions about the viability of seeding innovation from a single classroom, despite best efforts.
A. I. du Pont Middle and High Schools (Red Clay), A.I. du Pont Astronomy Cooperative Initiative Evaluation Total Score=21	This project provided repairs to and added to the district's astronomy equipment (planetarium and observatory). Additionally this project sought to increase student involvement in the astronomy clubs at both the middle and high school levels, fostering collaboration between the two. This project was not able to recruit the number of students they had hoped to the astronomy clubs. Two public presentations were held. Additionally this project provided insufficient evidence on student achievement outcomes. After consultation, the district was advised to revise its original proposal, resulting in a substantial financial savings. The district then reallocated those funds to the environmental sciences club to refurbish the greenhouse at A.I. du Pont Middle School. This represented a major change in strategy.
Gateway Charter School (Charter), Expansion of Reading Workshop Model Evaluation Total Score=18	This project purchased iPad minis, subscribed to an online book collection, and launched a "40 book challenge" in order to promote reading among its students. Participation was mixed: about half of Gateway students participated and a third completed the challenge. While evidence of DIBELS and DCAS growth for participating 3rd graders was provided, other grades did not show growth. While the plan called for implementing a Readers' Workshop model, only very minimal PD was provided and instructional change did not occur. Monitoring and dissemination were weaknesses. District support was not rated because Gateway is a charter school.
H.B. du Pont Middle School (Red Clay), C.E.L.E.B. (Cognitive and Emotional Learning and Esteem Building) Program Evaluation Total Score=17	This project provided a total of 15 sessions to identified students to support positive behavior, self-esteem, and attitudes towards school. Some of the associated activities in the proposal, such as outreach to parents, professional development for school staff, and field trips are not mentioned in the report, and tutoring for students occurred little if at all. Relationships between C.E.L.E.B. and the school staff are crucial to the sustainability of this effort; positive statements are made about this

	collaboration, but no evidence is provided. There was no decrease in disciplinary activity for participating students, and little other outcome evidence is provided. The late start and poor weather created challenges to implementation.
--	---

Individual SIIP project rubric scores are presented in Table 1 below. Four projects met or exceeded the “expected level” of performance across the seven process evaluation rubrics: BRINC, Red Clay 4-6th Grade Math Partnership Institute, Stanton Trauma-Informed System of Care, and Brandywine Biology. None of the 14 projects met the “expected level” for the five outcome evaluation rubrics. BRINC was the only one that met or exceeded the “expected performance” level for the total score. Red Clay 4-6th Grade Math Partnership Institute was the only project that scored a “2” or above on all of the rubrics and BRINC was the only project that scored a “5” on any rubric.

We next looked at the relative rankings of the individual projects to detect any patterns in the individual projects’ scores. The 14 projects were divided into two groups: those that scored above, at, or approaching expectation and those scoring below expectation for the process, outcome, and total rubric scores. Deciding where to draw the cut line was based on gaps in the distributions of scores. Table 2 summarizes these data.

There was some consistency across the top tier of projects. BRINC, Red Clay 4-6th Grade Math Partnership Institute, and Stanton Trauma-Informed System of Care scored in the top tier across both subsets of scores (process and outcome) and grand total. Brandywine Biology scored in the top tier for the process subset and for the grand total. Three other projects scored in the top tier for the process or outcome subset only; they included Gallaher GOALS, POLYTECH, and Cape Henlopen.

Table F1. SIIP Project Rubric Scores

Project	Process Rubrics								Outcome Rubrics						Goal Rubrics		Grand Total
	Project Plan	PD	Resources	Educator Participation	Student Participation	Monitoring	District Support	Total	Theory of Action	PD Application	Student Achievement	Future Plans	Dissemination	Total	Goal	Goal-Rating	
BRINC	3	3	4	4	4	3	5	26	1	2	1	5	4	13	4	1	40
RC Math	3	3	3	3	3	2	4	21	2	2	3	2	4	13	2	3	37
Stanton	2	3	4	3	3	4	2	21	2	1	2	3	4	12	3a	2	35
Brandywine Bio	3	3	3	3	4	2	4	22	2	2	2	3	1	10	2,4	2,3	34.5
Gallaher GOALS	4	3	4	2.5	2.5	2	2	20	3	2	2	2	1	10	4	2	32
POLYTECH	3	2.5	3	3	3	2	3	19.5	1	2	1	2	4	10	2	2	31.5
Cape Henlopen	2.5	2	3	3	3	1	2	16.5	2	2	2	2	4	12	4	2	30.5
Laurel i-Impact	2	2.5	2	3	2	1	3	15.5	2	1	1	2	2	8	4	1	24.5
RC SAT Prep	2	1	3	3	2	1	2	14	1	1	3	2	1	8	1,4	2,1	23.5
Warner WIZ	2	2	2	1	2	2	1	12	2	1	2	2	1	8	3a,4	2,3	22.5
Chipman History	4	N/A	3	2	3	2	1	15	1	N/A	2	1	1	5	1,2	3,1	22
RC Astronomy	2	2	2	3	1	2	2	14	1	1	1	2	1	6	2	1	21
Gateway	2	1	3	2	2	1	N/A	11	1	1	1	2	1	6	4	1	18
HB du Pont C.E.L.E.B.	2	1	1	2	2	1	1	10	1	1	2	1	1	6	3a	1	17
Mean	2.48	2.14	2.72	2.55	2.49	1.79	2.38	16.2	1.51	1.40	1.71	2.13	2.10	8.65	3.25	1.54	27.78
SD	0.74	0.98	0.86	0.72	0.84	0.86	1.38	4.70	0.65	0.63	0.70	0.97	1.46	2.76	1.70	0.95	7.44

Note: Rubric scores range from 1 (Minimal Performance) to 5 (Exceptional Performance). A score of 3 is the Expected Performance. An "N/A" score was assigned when the project did not include professional development (Chipman History) or the grantee did not have a district office (Gateway Charter).

Table F2. Rank Order of SIIP Projects Based on Process, Outcome, and Grand Total Rubric Scores

Score Band	7 Process Rubrics Expected Score=21	5 Outcome Rubrics Expected Score=15	13 Rubrics - Grand Total Expected Score=39
Above, At, or Approaching Expectation	• BRINC (26) • Brandywine Biology (22) • Red Clay Math Partnership (21) • Stanton Trauma-Informed Care (21) • Gallaher GOALS (20) • POLYTECH (19.5)	• Red Clay Math Partnership (13) • BRINC (13) • Cape Henlopen (12) • Stanton Trauma-Informed Care (12)	• BRINC (40) • Red Clay Math Partnership (37) • Stanton Trauma-Informed Care (35) • Brandywine Biology (34.5)
Below Expectation	• Cape Henlopen (16.5) • Laurel i-Impact (15.5) • Chipman History (15) • Red Clay SAT Prep (14) • Red Clay Astronomy (14) • Warner WIZ (12) • Gateway (11) • HB du Pont C.E.L.E.B. (10)	• Gallaher GOALS (10) • POLYTECH (10) • Brandywine Biology (10) • Laurel i-Impact (8) • Red Clay SAT Prep (8) • Warner WIZ (8) • Gateway (6) • HB du Pont C.E.L.E.B. (6) • Red Clay Astronomy (6) • Chipman History (6)	• Gallaher GOALS (32) • POLYTECH (31.5) • Cape Henlopen (30.5) • Laurel i-Impact (24.5) • Red Clay SAT Prep (23.5) • Warner WIZ (22.5) • Chipman History (22) • Red Clay Astronomy (21) • Gateway (18) • HB du Pont C.E.L.E.B. (17)

Note: Rubric scores range from 1 (Minimal Performance) to 5 (Exceptional Performance). A score of 3 is the Expected Performance. Rubrics #13-16 focused on the attainment of project-specific goals. Most projects were only rated on one of these rubrics. A few projects identified more than one goal; in these cases, the rubric scores were averaged.

Cross-SIIP Project Findings

We next looked at the individual rubric scores across all of the projects. Table F3 summarizes the frequency distribution of scores for the process, outcome, and goal-specific evaluation rubrics.

Table F3. Distribution of Individual Rubric Scores

Rubric Scores	Process		Outcome		Goal	
	n	%	n	%	n	%
5 Exceptional Performance	1	1.0	1	1.4	0	0.0
4	11	11.5	5	7.2	0	0.0
3 Expected Performance	32	33.3	5	7.2	2	11.1
2.5	5	5.2	0	0.0	0	0.0
2	33	34.4	29	42.0	7	38.9
1 Minimal Performance	14	14.6	29	42.0	9	50.9

The 14 SIIP projects generally scored higher on the process outcome rubrics than on the outcome or goal-specific rubrics. Almost half of the individual process rubric ratings (45.8%) were at the “expected performance” level in contrast to a third of the individual outcome ratings (33.3%) or about a tenth of the goal-specific ratings (11.1%). In our experience, this is not unexpected during the first year of operation for any program in which program staff efforts are focused on implementing a new program. In addition, it is often difficult to achieve intended outcomes in the first year because of the newness of the project and the time necessary to demonstrate results.

We next looked to see if the 14 SIIP projects scored higher on some rubrics than on others. Table F4 summarizes these data. Consistent with the results in Table 3 above, individual process evaluation rubrics generally obtained higher mean scores than individual outcome evaluation rubrics.

Table F4. Distribution of Scores by Rubric

Rubric		Rubric Score														Mean	SD
		1		2		2.5		3		4		5		N/A			
		n	%	n	%	n	%	n	%	n	%	n	%	n	%		
Process Rubrics	Project Plan	0	0	7	50	1	8	4	29	2	14	0	0	0	0	2.48	.74
	PD	3	21	3	21	2	14	5	36	0	0	0	0	1	8	2.14	.98
	Resources	1	8	3	21	0	0	7	50	3	21	0	0	0	0	2.72	.86
	Educator Participation	1	8	3	21	1	8	8	57	1	8	0	0	0	0	2.55	.72
	Student Participation	1	8	5	36	1	8	5	36	2	14	0	0	0	0	2.49	.84
	Monitoring	5	36	7	50	0	0	1	8	1	8	0	0	0	0	1.79	.86
	District Support	3	21	5	36	0	0	2	14	2	14	1	8	0	0	2.38	1.38
Outcome Rubrics	Theory of Action	7	50	6	43	0	0	1	8	0	0	0	0	0	0	1.51	.65
	PD Application	7	50	6	43	0	0	0	0	0	0	0	0	1	8	1.40	.65
	Student Achievement	5	36	7	50	0	0	2	14	0	0	0	0	0	0	1.71	.70
	Future Plans	2	14	9	64	0	0	2	14	1	8	0	0	0	0	2.13	.97
	Dissemination	8	57	1	8	0	0	0	0	4	29	1	8	0	0	2.10	1.46

Note: Rubric scores range from 1 (Minimal Performance) to 5 (Exceptional Performance). A score of 3 is the Expected Performance.

Highest mean scores were obtained for four process rubrics: resources (Mean=2.72), recruitment and participation of educators and students (Means =2.55 and 2.49 respectively), and project plans (Mean=2.48). Many of these grants brought funds to the district or school that were essential to project implementation and so it is not surprising that technology, supplies, and materials were purchased and used as expected. The majority of projects required teachers or students in particular grades to participate; recruitment was only an issue in the Red Clay SAT Prep and Astronomy projects. Most project plans were originally developed for only one year of funding and were generally followed; modifications were made to accommodate delays in the purchasing of technology or other supplies or inclement weather.

Both the Professional Development and Monitoring mean scores were the two lowest process evaluation outcome rubrics. Three projects (i.e., Red Clay SAT Prep, Gateway, and H.B. du Pont C.E.L.E.B.) did not deliver the professional development intended and five others (i.e., POLYTECH, Cape Henlopen, Laurel i-Impact, Warner WIZ, and Red Clay SAT Prep) did not report evidence of teacher implementation of the professional development knowledge or skills or did not provide follow-up support. Five of the 14 projects (i.e., Cape Henlopen, Laurel i-Impact, Red Clay SAT Prep, Warner WIZ, Gateway, and H.B. du Pont C.E.L.E.B.) provided only minimal evidence of any project monitoring. Six others (i.e., Brandywine Biology, Chipman History, GOALS, POLYTECH, Red Clay Astronomy, Red Clay 4-6th Grade Math Partnership Institute) scored low because they did not indicate who did the monitoring or what data monitors examined.

The two highest outcome evaluation means were Future Plans (Mean=2.13) and Dissemination (Mean=2.10). All but two of the projects scored at least a "2" indicating that school or district had developed plans to address ongoing needs, though these plans were incomplete or did not seem entirely feasible. More than any other rubric, Dissemination scores were bimodal with almost all scores at either a "4" or a "1". The former indicates that one-way dissemination has begun to other schools and districts while the latter means that few plans have been developed. This rubric stresses dissemination external to the district (or charter school); eight of the 14 projects had shared findings within their district but had not yet taken steps to communicate with other districts. Dissemination activities often are scheduled to occur at the end of a project and so these scores may be conservative estimates of what districts and schools eventually share.

Rubric mean scores were lowest related to the Application of Professional Development (Mean=1.40), Articulated Theory of Action (1.51), and Student Achievement (Mean=1.71). The Application of Professional Development scores were low. Seven projects provided no evidence that teachers applied knowledge and skills gained during professional development in their classrooms; six others provided incomplete evidence and/or evidence that indicated inconsistent application of professional development. In the latter case, this was often an issue of manpower to observe sufficient numbers of teachers in the classroom. In two cases (i.e., Laurel i-Impact and Gateway), teachers received technology later than expected and so implementation was limited. Relatively low performance on the second rubric, Articulated Theory of Action, is most likely an artifact of the reporting guidelines. Most projects had presented a theory of action in the proposal documents submitted to the DE DoE and simply did not repeat it in the final report. Evidence to support student achievement gains was missing

or incomplete for 12 of the projects; only the Red Clay 4-6th Grade Math Partnership Institute and Red Clay SAT Prep scored at the “expected performance” level. In the former, professional development was targeted at very specific math skills tested by DCAS and the project was able to demonstrate improvement on those skills. The Red Clay SAT Prep demonstrated higher scores on students’ SAT scores than predicted by their PSATs.

The DE DoE originally proposed four goals for SIIP. Some of the potential grantees were initially confused and thought that they had to respond to all four. As a result, we asked them to clarify special goals that their SIIP projects addressed. The most frequently selected goals were Accelerating the Achievement of Under-Performing Groups of Students (n=8) and Common Core Implementation and Assessment (n=5). Three projects intended to improve School Climate for Students and two focused on Teacher-Led Projects. Four projects addressed two goals three of the four were focused on Accelerating the Achievement of Under-Performing Groups of Students with one of the other three goals (see Table F1 for a complete breakdown). Table F5 summarizes the distribution of rubric scores by goal.

Table F5. Distribution of Rubric Scores by Project Goal

Goal	Rubric Score									
	1		2		3		4		5	
	n	%	n	%	n	%	n	%	n	%
Teacher-Led Projects (n=2)	0	0.0	1	5.6	1	5.6	0	0.0	0	0.0
Common Core Implementation and Assessment(n=5)	2	11.1	2	11.1	1	5.6	0	0.0	0	0.0
School Climate – Students (n=3)	1	5.6	2	11.1	0	0.0	0	0.0	0	0.0
School Climate – Teachers (n=0)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Accelerating the Achievement of Under-Performing Groups of Students (n=8)	4	22.2	2	11.1	2	11.1	0	0.0	0	0.0

Note: Percentages are based on the number that obtained each rubric score divided by the total number of goals addressed (n=18).

Four of the projects scored at the “expected performance” level for their project-specific goals. The Red Clay 4-6th Grade Math Partnership earned “3s” for Common Core Implementation and Assessment, Brandywine Biology and Warner WIZ for Accelerating the Achievement of Under-Performing Groups of Students, and Chipman History for Teacher Leadership. Five other projects earned a score of “2” including Stanton Trauma-Informed Care, Gallaher GOALS, POLYTECH, Cape Henlopen, and Red Clay SAT Prep. Many of the scores are depressed because of the limited or incomplete data presented in final reports. The Accelerating the Achievement of Under-Performing Groups of Students goal proved to be the most challenging, most likely because of the relatively short time in which to improve student performance.

Conclusions and Recommendations

The SIIP grant program provided an opportunity for the DE DoE and the DERDC to collaborate on the evaluation of the 14 projects. The individual grantees were required to conduct their own evaluations as part of their funding requirements. DERDC developed a set of process and outcome evaluation rubrics that were used to assess the implementation of the individual projects as well as their attainment of project outcomes and common goals. The findings of this metaevaluation are relevant to assessing the overall success of the SIIP program as well as to future state funding initiatives.

Overall, SIIP set ambitious goals for the 14 grantees. Only one of the projects reached the “expected performance” level overall. Almost half of the projects scored at the “expected performance” level for the process evaluation rubrics and about a third on the outcome rubrics. A number of factors help explain these results, some related to the design of the SIIP program itself. The first order of business for any new grantee is to “get their projects up and running.” The SIIP projects followed this maxim and prioritized project operations. As a result, projects consistently scored the highest on the process rubrics, particularly those related to having a project plan, using resources, and recruiting and engaging teachers and students.

The projects varied tremendously in terms of scope and focus. Although the most expansive project also was the highest scorer (i.e., BRINC), this project also had resources and active support from four districts that likely contributed to its success. Many of the other projects had minimal resources beyond those paid for by the grant and limited active engagement and support from others in the district. Severe winter weather also challenged many projects to conduct professional development and other activities as originally planned. In addition, the timeline available to demonstrate improvements in the five project-specific goals was comparatively short (e.g., Common Core State Standards implementation and assessment, Closing the Achievement Gap).

Two other factors (not directly tied to the design of the SIIP program) likely contributed to the above results. First, the skills needed to write winning proposals and operate sound educational programs are not the same skills needed to conduct evaluations or draft final reports. In spite of several offers of support, few grantees asked for assistance throughout the year and none asked for help on preparing their final report.⁴ Although the final reports, as a group, were much stronger than the mid-year reports submitted to the DE DoE, many were incomplete and lacking in evidence to support claims. In addition, grantees knew that additional years of funding would not be forthcoming and so likely did not invest significant time in preparing their reports. These factors most likely contributed to depressed scores on the rubrics.

Three recommendations are offered to help guide future DE DoE initiatives. These recommendations stem from the metaevaluation findings discussed above.

⁴ We helped four projects with instrumentation, reviewed and gave feedback on midyear reports to two projects, observed activities for two projects, and assisted one project with data analysis. Overall, five of the 14 received assistance from DERDC.

1. Limit scope of work for one-year projects.

Ideally, educational projects should rarely be funded for a single year. However, revenue streams available to state departments of education often are limited to a single year. They are thus caught in an impossible quagmire. Rather than not fund any projects, our advice is to limit such projects severely in terms of scope and focus. Set goals that can realistically be attained in a single year.

2. Set explicit expectations in the Request for Proposals (RFPs).

Use the RFP to communicate what grantees are expected to do. Be specific about required activities, particularly if there are design features that are important to the DE DoE. For example, professional development without follow-up support and monitoring is unlikely to change instructional practice. If the funding is going to support professional development, it may be helpful to specify those two features in the RFP. Based on this recent round of funding, we would recommend more explicit expectations for not only professional development, but also for project monitoring and evaluation procedures and reporting of evaluation findings.

3. Continue to develop and use the evaluation rubrics to evaluate DE DoE initiatives.

The development and use of the evaluation rubrics with the SIIP program provides a unique opportunity to look at the success of one DE DoE initiative. Up to the present, DE DoE has not systemically looked at such initiatives, either to determine their relative success or what lessons could be learned about funding future initiatives. State departments of education have limited discretionary funds at their disposal. In order to maximize their potential benefit, valid and reliable data are needed about their performance on a constant set of indicators. This set of rubrics, with some modifications, provides DE DoE with the possibility of creating such a dataset. Use of the evaluation rubrics also would send an important message to grantees about the state's expectations that accompany such funding.

SIIP Evaluation Rubrics

The SIIP evaluation rubric will be used to assess the performance of SIIP grant projects, where “1” represents minimal performance, “3” represents expected performance, and “5” represents exceptional performance towards grant project plan activities and goals.

CRITERION	PERFORMANCE RATINGS – PROCESS EVALUATION CRITERIA						EVIDENCE
	N/A	1 Minimal Performance	2	3 Expected Performance	4	5 Exceptional Performance	
Grant Project Plan		Minimal number of project activities are carried out. Incomplete or vague justification is provided for abandoning original plans.	Some, but not all, project activities are carried out. Justifications for changes are incomplete.	Most project activities are carried out. Justifications are provided when changes are made.	Most project activities are carried out. Opportunities are identified to extend project beyond original scope (e.g. through partnerships, linkage of programs).	Most project activities are carried out. Opportunities to extend project beyond original scope are acted on.	
Professional Development (PD)		Minimal PD is delivered.	Some PD is delivered, but does not reach entire targeted audience. Minimal implementation of PD knowledge and skills occurs.	PD is delivered to targeted audience. PD knowledge and skills are generally implemented as intended.	PD is delivered to targeted audience and generally implemented as intended. Some additional follow-up support occurs (e.g., coaching, monitoring, peer collaboration).	PD is delivered to targeted audience and generally implemented as intended. Consistent follow-up support occurs.	
Resources (Technology/ Capital/ Tangible Supplies and Materials)		Minimal resources are purchased or used.	Resources are used, but not always for intended purposes.	Resources are used as intended.	Resources are used as intended. Plans exist to extend their reach to support other related activities.	Resources are used as intended. Project resources extended to support other related activities.	

CRITERION	PERFORMANCE RATINGS – PROCESS EVALUATION CRITERIA						EVIDENCE
	N/A	1 Minimal Performance	2	3 Expected Performance	4	5 Exceptional Performance	
Recruitment and Participation (Educators)		Targeted population is not recruited or does not participate.	Targeted population is recruited but sufficient numbers do not participate.	Targeted population is recruited and sufficient numbers participate.	Targeted population is recruited and sufficient numbers participate. Varying levels of engagement occur.	Targeted population is recruited and sufficient numbers participate. High levels of engagement occur.	
Recruitment and Participation (Students)		Targeted population is not recruited or does not participate.	Targeted population is recruited but sufficient numbers do not participate.	Targeted population is recruited and sufficient numbers participate.	Targeted population is recruited and sufficient numbers participate. Varying levels of engagement occur.	Targeted population is recruited and sufficient numbers participate. High levels of engagement occur.	
Monitoring		Minimal monitoring of grant activities occurs.	Inconsistent and/or incomplete monitoring of grant activities occurs.	Consistent and complete monitoring of grant activities occurs.	Consistent and complete monitoring of grant activities occurs. Monitoring is sometimes used to inform practice.	Consistent and complete monitoring of grant activities occurs. Monitoring is consistently used to inform practice.	
District Support		District offers minimal support of grant project activities.	District offers some support for grant project activities, but it limited or inconsistent.	District supports grant project activities (e.g., staffing support, professional development).	District is actively engaged in all grant project activities.	District is actively engaged in all grant project activities, advocates for the project's success.	

CRITERION	PERFORMANCE RATING –OUTCOME EVALUATION CRITERIA – ALL GOALS						EVIDENCE
	N/A	1 Minimal Performance	2	3 Expected Performance	4	5 Exceptional Performance	
Articulated Theory of Action		Minimally articulated theory of action with few ties to targeted population needs.	Theory of action is articulated but lacks clarity. Some ties exist to targeted population needs.	Clear theory of action is articulated. Clear ties exist to targeted population needs.	Clear theory of action is articulated. Theory has some evidence to support it. Clear ties exist to targeted population needs.	Clear theory of action is articulated. Theory is grounded in empirical research. Explicit ties to targeted population needs.	
Application of Professional Development (PD)		Few participating educators incorporate targeted instructional practices.	Some participating educators incorporate targeted instructional practices.	Most participating educators incorporate targeted instructional practices.	Most participating educators incorporate targeted instructional practices with mixed fidelity.	Most participating educators incorporate targeted instructional practices with high fidelity.	
Student Achievement		Achievement targets are not met.	Some achievement targets are met.	Most achievement targets are met.	Most achievement targets are met and some exceed targets.	Most achievement targets are met and many exceed targets.	
Future Plans to Address Ongoing Needs		Few plans are developed to address ongoing needs.	Plans to address ongoing needs are developed but are incomplete or not feasible.	Plans to address ongoing needs are developed and feasible.	Plans to address ongoing needs are beginning to be implemented.	Plans to address ongoing needs are implemented.	
Dissemination of Information		Few plans are developed to share information to other schools and districts.	Plans are developed to share information to other schools and districts, but they are incomplete or vague.	Plans to share information to other schools and districts include details (e.g., data) and materials have been developed.	One-way dissemination of information to other schools and districts has begun.	Two-way dissemination of information to other schools and districts has begun.	

CRITERION	PERFORMANCE RATING – OUTCOME EVALUATION CRITERIA – SPECIFIC GOALS						EVIDENCE
	N/A	1 Minimal Performance	2	3 Expected Performance	4	5 Exceptional Performance	
Goal #1 – Teacher Leadership		Teachers' roles in project is limited and/or pro forma.	Teachers' roles in project is substantive. Teachers are part of the decision-making processes.	Teachers play significant leadership roles in project (e.g., makes decisions or sets direction).	Teacher leadership begins to spread (e.g., non-project teachers begin to take on leadership for instructional improvement and student achievement).	Teacher leadership spreads significantly.	
Goal #2 - Common Core		Minimal implementation of aligned curriculum to CCSS as planned.	Partial implementation of aligned curriculum to CCSS as planned.	Substantial implementation of aligned curriculum to CCSS as planned. Fidelity varies greatly.	Substantial implementation of aligned curriculum to CCSS as planned. Moderate level of fidelity found across all targeted classrooms.	Substantial implementation of aligned curriculum to CCSS as planned. High levels of fidelity found across all targeted classrooms.	
Goal #3a – School Climate (Students)		Few school climate targets/goals are met.	Some school climate targets/goals for students are met.	Most school climate targets/goals for students are met.	Most school climate targets/goals for students are met and some exceed targets.	Most school climate targets/goals for students are met and many exceed targets.	
Goal #3b – School Climate (Teachers)		Few school climate targets/ goals are met.	Some school climate targets/ goals for teachers are met.	Most school climate targets/ goals for teachers are met.	Most school climate targets/ goals for teachers are met and some exceed targets.	Most school climate targets/ goals for teachers are met and many exceed targets.	

CRITERION	PERFORMANCE RATING – OUTCOME EVALUATION CRITERIA – SPECIFIC GOALS						EVIDENCE
	N/A	1 Minimal Performance	2	3 Expected Performance	4	5 Exceptional Performance	
Goal #4 – Under-achieving Student Groups		Targeted achievement gap has narrowed minimally if at all.	Targeted achievement gap has narrowed somewhat, but the closing is not substantial.	Targeted achievement gap has narrowed.	Targeted achievement gap has narrowed significantly.	Targeted achievement gap has been closed.	
Project-Specific Outcomes (please specify below)		Achievement targets are not met.	Some achievement targets are met.	Most achievement targets are met.	Most achievement targets are met and some exceed targets.	Most achievement targets are met and many exceed targets.	

Appendix G

ARTIFACT G: REFLECTIONS FROM TEACHING EDUC 774, SPRING 2016

Introduction

After the seven-week whirlwind that was Designing Professional Development (EDUC 774), this essay gives me an opportunity to reflect on this new course, identifying strengths as well as issues. I also reflect more broadly on what I gained through the preceptor experience. I describe lessons learned – and questions developed – about instruction in an online masters’ program in teacher leadership. Finally, I discuss how the course influenced my thinking about professional development (PD).

Reflections on EDUC 774

This was the first time this course was offered, and the first course in the Teacher Leadership program to be delivered through UD’s partnership with Wiley. I joined the team after the instructor, Dr. Chrystalla Mouza, completed its design.

Our 58 students had varied backgrounds. The largest group (22, 37.9%) were elementary school teachers; another 11 (19%) were middle school teachers and the same number (11, 19%) were high school teachers. Eleven students worked in higher education as instructors (8, 13.8%) or administrators (3, 5.2%). One middle school assistant principal and one professional development (PD) provider rounded out the group. Some students were mid-career; others had only a few years of teaching experience. Because of the structure of the teacher leadership program, we had new students, students in their penultimate (ninth) class, and everything in between. We

also had four doctoral (EdD) students, taking this as an elective, and one masters student in educational technology.

By several measures, the course was a success. It had a large enrollment, and all 58 students completed all course requirements. Moreover, almost all students kept pace – essential in such a short class. There were 12 assignments; for three of them (25%), 58/58 students submitted on time. For six assignments (50%) one or two students submitted late (> 5%), almost always after requesting permission. For two assignments including the first, four or more students (>5%) submitted late.

Student performance in the course was also strong. The distribution of final grades was: 41 A (71.7%), 16 A- (27.6%) and 1 B+ (1.7%). Since assignments were structured similarly throughout the course, the students who struggled with their early discussion posts or PDP entries generally improved as the semester went on. Four students required individual support and/or intervention with in-person or telephone meetings. Overall most students demonstrated a good understanding of course concepts and invested sufficient time in their work.

Our student evaluation data is presented in Appendix 1. Evaluation questions were phrased in terms of “the instructor” and “the course.” Dr. Mouza and I divided the students in half for grading and commenting, switching groups every week. Some students used plural pronouns (“they”) suggesting that they were responding about both of us. Others referred to Dr. Mouza by name. Even though the evaluation did not ask specifically about “the preceptor,” and even though I was not directly involved in the course design, I still found the results meaningful because they communicated what students valued about the class, and what needs improvement.

Overall feedback was positive. By large margins, students found EDUC 774 relevant and well organized, and found the instructors responsive and helpful. In such an accelerated class, a quick turnaround is important. We received high marks for answering questions promptly and giving meaningful feedback, but several students complained about the lag time in grading.

This course was very informative and very well put together. The professors responded very quickly with explicit directions on how to complete tasks. So far, it's the best course I've taken in the program!

Grades were not posted in a timely enough manner. Most of the class were completing the next week's assignment before receiving the last weeks grades, so any feedback received can not be incorporated until 2 weeks later.

With 58 students and two assignments to be graded each week, the load was heavy, even with two instructors working at what they felt to be a reasonable pace. Adjusting the expectations (e.g., not grading as many assignments, using peer grading) may be required. Several students wrote that the rubrics needed to be more detailed, a suggestion that might also speed up grading. Wiley also provided some suggestions for large classes that may be relevant, such as developing group rather than individual assignments (Burchfiel, 2016).

Organization and transparency are of paramount importance in an online, asynchronous learning environment (Anderson, Rourke, Garrison & Archer, 2001). This was especially crucial for us since all modules were unlocked. Students appreciated having full access from day 1 but it was more work up-front for Dr. Mouza. In this situation, mid-course corrections or clarifications are difficult. As recommended by Burchfiel (2016), Guhlin (2009) and Johnson (2013), Dr. Mouza and I checked the course carefully before it went live to minimize problems. Students

generally found the course easy to navigate, with a manageable workload. Several stated that they appreciated how the professional development plan (PDP) was chunked across the whole seven weeks. The PDP template provided additional scaffolding for this project.

The course was well organized and the information presented was clear and useful.

I love how the final project was scaffolded in each week's module. It made the project flow so much better and I was able to digest each portion prior to moving to the next.

The predictability of the course had its drawbacks, however. Some students found the structure, with a discussion post and a PDP section due most weeks, repetitive, or stated that assignments felt redundant. This is the result of the scaffolding praised above. The presentations in Module 4/5 were an opportunity for students to demonstrate different skills, giving them – and us – a welcome break from discussion boards. The challenge for the future would be to incorporate more variety in assignments without making the course too complicated. Relatedly, some students felt that discussion posts were trying to address too many objectives – foster discussion, ensure that students read, assess students' understanding of concepts, etc. Expectations for these posts could be reduced, but only if other assignments were added, which might make the course cumbersome. These student comments highlight a set of trade-offs to be considered.

As mentioned above, most students had positive comments about our accessibility and helpfulness as instructors. This was affirming to me, as one of my roles was to look for student questions, answering them if I could and passing them on if not. Some students reached out for extra help. Yet at least one student felt s/he did not have a strong connection with us:

Of all the courses I've taken so far in this program, I felt the most disconnected with the instructors in this one. I'm not exactly sure why - perhaps the program is growing beyond the capacity of the instructors? In past courses, the instructors used informal videos and personal messages to communicate with the class thru Monday Morning messages and I felt like I got to know the instructors better through those strategies. Overall, I feel that the content was sufficient, though the activities were not as engaging as in past courses.

This student provides some suggestions for improving our online presence, and echoes Guhlin's (2009) suggestion to use multimedia to personalize a class. Another way to do this is through video or audio feedback on assignments. Other authors suggest communicating pro-actively with students through brief check in emails (Johnson, 2013) or "mass direct emails" to segments of the class, such as students who improved from one assignment to the next (Berry, 2009).

Personal Reflections

This independent study provided me with a totally different teaching experience from any I had had before – different in scale, format, audience, and topic.

Scale & Format: I taught middle and high school English in independent schools where my classes had 12 – 15 students, gathered around a Harkness table. I then ran an educational out-of-school time program where we capped classes at eight. Later, I coached graduate research assistants at DERDC and, most recently, taught 2nd grade RTI and 4th grade accelerated reading groups at my daughters' school. In all these settings, small equaled good, allowing for individual attention, differentiation, and long responses to student work. And relationships were built through extensive, face-to-face time together. So adjusting to an online environment with 58 students posed a new challenge.

As Dr. Mouza suggested to me mid-semester, I tended to over-invest in commenting on student work, and can fall into “The Teaching Trap” (Rockquemore, 2016). My efficiency increased over the seven weeks, but I have a ways to go. My visceral reaction to some of the recommended strategies (e.g., Dr. Fred’s master list of feedback phrases; Johnson (2013)’s suggestion that only the bottom 20% of students need personal comments) was resistance. I think that comes from my prior experiences, outlined above. But I need to weigh whether faster, more systematized feedback is more useful to my students than more detailed and individual comments that take longer. Even though I am not on a tenure track, I do have other responsibilities besides teaching, and there is a danger of sacrificing longer-term projects for more urgent tasks like grading.

I also need to update my expectations and strategies for getting to know students. The comment about disconnection, cited above, bothered me. It was hard for me to retain information about our many students, except for those I knew personally. Next year, I want to work harder to build rapport up front and develop systems for “remembering” students. These steps don’t have to take a lot of time: quick check-in emails, teaching journals, more use of technology tools to hear students’ voices and see their faces, even just asking students to update their Canvas profile so I can click and see who they are. It might also help if we have some repeat students in the fall.

A more intellectual way to forge connections is “facilitating discourse” (Anderson et al, 2001). I would like to do more of this, with a different assignment format. This spring we did not have much time to participate in the discussion boards as we were always assessing them. As a face-to-face teacher, I am skilled at facilitation – encouraging students to talk to and question each other, posing

occasional critical questions, highlighting points of convergence and contrast in the dialogue. I would like to get better at doing these things online.

Audience & Topic: Obviously, teaching professional development (PD) to teacher leaders was very different from teaching English or reading to 2nd – 12th graders. I had some background teaching doctoral students about PD, but the focus was specifically on evaluation and we were working one-on-one. Like all adult learners, our EDUC 774 students brought beliefs based on prior experience with them. Many of these beliefs were strong; many of these PD experiences were negative (Davis & Andrzejewski, 2009). Not surprisingly, our students had often been subjected to “sit and get,” ineffective PD. Some had also seen bright spots of PD that *did* align with the research.

Whether they were working from best practices or “what not to do,” students generally embraced the idea that PD can – and needs to – improve, and that it matters. All students were motivated on one level; EDUC 774 is a requirement for the degree that almost all of them were pursuing. Many also stated individual motivations and specific applications for the work. These were especially frequent among students who had already attained some level of leadership, as in these comments from the introductory post:

I recently wrote an ELA/SS curriculum and am involved in planning the PD to introduce it to the other K teachers in the district. I have found that the 30+ hours writing the curriculum was FAR LESS stressful than sitting in a room with principals and supervisors to PLAN the PD for it. I hope that this course will give me the confidence and skill set to voice my opinion more during these meetings.

Currently [I] am on my way toward becoming a lead teacher / instructional coach. I hope that my participation in this course will help me to become a better, more well-informed lead teacher and coach next year in my 7th grade position. I will be taking on a brand new

staff...and hope to plan professional development that motivates my new team to be the best they can be. I have worked closely with the admin team at my school to plan professional development on a small scale, but have never learned "how" to implement effective professional development.

As described earlier, our students varied in their roles and seniority. For some, the opportunity to plan, deliver, or influence PD in their schools was still on the horizon. These students expressed more general ideas about what they hoped to gain from the course, or how they would apply the ideas. Despite this range, many of our students demonstrated all five of Knowles' (1980, 1984) characteristics of adult learners in their self-concept, experience, readiness, orientation, and motivation to learn. It was exciting to see one of the principles we were teaching about reflected in our own class. As a secondary ELA teacher, I sometimes had to convince my students why they should care about what we were learning, or how they would use it. This was never a problem in EDUC 774.

I found some aspects of our topic and audience more challenging. Perhaps because it was my first time in this course, or because I never "learned" PD or teacher leadership in this way, I did not always have a clear picture of what we wanted our students to know and be able to do, and how we would know. More explicit rubrics probably would have helped me, too! It was evident enough whether students understood the readings and resources, but the quality of application was harder to discern. Some PD plans were more authentic than others, but was this because of student skill, effort and thoughtfulness, or because their context fit the assignment better? Some plans were more feasible than others, but this often reflected differences in student role/responsibility. Some had a better understanding of what it takes to actually implement any plan. I tried not to penalize students if their roles (e.g., newer teacher, higher education administrator) were less conducive to the task.

Another challenge was to decide where the bar should be in such a compressed course. For example, based on my background in evaluating PD I could see that most students only had a superficial understanding of this topic (Module 7). But it was the last week of the course; they had only had a few days to learn and think about evaluation. What could we reasonably expect?

As I reflected on the course, I looked back at our syllabus and found these Teacher Leader model standards, to which it aligned:

DIID: Teach and support colleagues to collect, analyze, and communicate data from their classroom to improve teaching and learning.

DIIB: Use information about adult learning to respond to the diverse learning needs of colleagues by identifying, promoting and differentiated professional learning.

DVD: Work with colleagues to use assessment and data findings to promote changes in instructional practices or organizational structures to improve student learning.

DIID. Identify and uses appropriate technologies to promote collaborative and differentiated professional learning.

DIIC. Facilitate professional learning among colleagues.

(Teacher Leadership Exploratory Consortium, 2012)

Revisiting these standards more often during the course might have helped me, and the students, feel more grounded. I think most of our work focused on DIIB, with some attention to DIID in Module 6. Although we asked students to provide data to justify their PDPs, we did not specifically prepare or support them for data analysis to drive PD planning (DIID and DVD). Our course included concepts to consider in facilitating PD (DIIC) but actually facilitating a PD session fell beyond the scope of our course. It might be interesting to follow up with our students midyear next year to

see whether they implemented their plans, and with what results. Finally, teaching this course gave me a chance to learn how *teachers* think about PD. In previous positions, I had more direct contact with how administrators, evaluators, and program managers view PD. When I worked at DASL, I provided PD for school leaders, which often entailed talking with them about PD for their staff members. I still do one project with DASL, Comprehensive Success Reviews (CSR) for schools under improvement. CSRs often include analysis of a school's PD, and recommendations for strengthening it. At DERDC, I evaluated a wide range of PD programs. While I collected data from teachers (asking how satisfied they were with the PD, what they learned from it, etc.), I generally was evaluating the implementation and effectiveness of a PD initiative or grant and taking a more macro view. In other words, I did not usually consult teachers in an open-ended way about their PD needs, hopes, and concerns. The readings in EDUC 774, especially in the first few modules, emphasized giving teachers voice and choice in PD. Differentiation, personalization, flexibility, leveraging teachers as facilitators and experts – all were encouraged. Discussion board posts and PDPs show that our students embraced and applied these “teacher-centered” ideas.

It was affirming to see students responding so positively to some of the PD and adult learning research. At the same time, I heard other voices in my head, questions that might have been asked in a CSR meeting or in an administrator-only PD session. *Do teachers know what areas they need to improve? If every staff member chooses PD individually, how do we make sure the new ideas cohere with each other, and align with school priorities? When PD is decentralized, job-embedded and differentiated, how do we monitor its quality and transfer to the classroom? When is all-staff PD or bringing in external expertise appropriate?*

A few of our students had administrative or school wide responsibilities, and I think they tended to be more able to think in bigger-picture terms about PD. Many students seemed to have a narrower frame of reference, focused very much on their individual experiences and classrooms. This was also evident in students' struggles to understand the material on PD policies and leadership in Modules 1 and 2. In general, I would have liked more opportunities to pose questions like these to challenge students' thinking.

One aspect of teacher leadership is perspective taking – the ability to look at a school/educational situation through others' eyes. This competency is directly addressed in Teacher Leader Model Standard DIB (“Models effective skills in listening, presenting ideas, leading discussions, clarifying, mediating, and identifying the needs of self and others in order to advance shared goals and professional learning”) and is implicit in most of the other standards (Teacher Leadership Exploratory Consortium, 2012). In the future, we might try to push our students to take various perspectives on their PDPs. What might a less motivated colleague say? What about a veteran teacher in his/her final year before retirement? How might a school administrator or board member respond? This final question is particularly important, considering that most of our students would require administrative support to bring their plans to fruition. Some of our students took the initiative to share their PDPs with their supervisors and/or colleagues or reflected on how they might do this. I suggest we require this step in the future as part of the final PDP submission. For instance, students could develop a presentation or memo.

Working as a preceptor in this course was a rich learning experience, and I am grateful for the opportunity. I hope to further investigate some of the reflections shared

here through a content analysis of student PDPs, which might eventually become an artifact for my own ELP.

References

- Anderson, T., Rourke, L., Garrison, D.R. & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks*, 5 (2), 1 – 17.
- Berry, R.W. (2009). Meeting the challenges of teaching large online classes: Shifting to a learner focus. *Journal of Online Learning and Teaching*, 5 (1). Retrieved May 10, 2016 from http://jolt.merlot.org/vol5no1/berry_0309.htm
- Burchfiel, G. (2016). Teaching and designing for large online classes. [PowerPoint slides].
- Davis, H. & Andrezejewski, C. (2009). Teacher beliefs. Retrieved May 10, 2016 from <http://www.education.com/reference/article/teacher-beliefs/>
- Guhlin, M. (2009). Eight tips for successful online course facilitation. Retrieved May 10, 2016 from http://www.educationworld.com/a_tech/columnists/guhlin/guhlin014.shtml
- Johnson, A. (2013). *Excellent Online Teaching*. Denver, CO: Aaron Johnson.
- Knowles, M. (1980, 1984). The adult learning theory – andragogy – infographic. Retrieved May 10, 2016 from <http://elearninginfographics.com/adult-learning-theory-andragogy-infographic/>
- Rockquomore, K.A. (2016). The teaching trap. [Online newsletter article]. National Center for Faculty Development and Diversity.
- Teacher Leadership Exploratory Consortium (2012). Teacher leader model standards. Retrieved May 10, 2016 from https://www.ets.org/s/education_topics/teaching_quality/pdf/teacher_leader_model_standards.pdf

Appendix 1
Summary of Student Evaluation Data for EDUC 774, Spring 2016

Table 1
Ratings of instructor in EDUC 774

	Strongly agree	Agree	Neither	Disagree	Strongly disagree
The instructor responded to my questions in a timely manner.	18 (56.2%)	12 (37.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
The instructor graded assignments in a timely manner.	16 (50.0%)	11 (34.4%)	2 (6.2%)	3 (9.4%)	0 (0.0%)
Rate your satisfaction with the quality of the instructor's feedback.	12 (37.5%)	17 (53.1%)	1 (3.1%)	2 (6.2%)	0 (0.0%)
My instructor was actively engaged in the course.	15 (46.9%)	14 (43.8%)	2 (6.2%)	1 (3.1%)	0 (0.0%)
The instructor cared about my learning progress.	17 (53.1%)	11 (34.4%)	3 (9.4%)	1 (3.1%)	0 (0.0%)

N=32

Table 2
Comments about instructor in EDUC 774

Response	Number	Percentage
Positive comments about instructor feedback or helpfulness	10	66.7
Positive comments about organization of course	6	40.0
Positive comments about content or utility of course	4	26.6
Critiques about organization of course	4	26.6
Critiques about timeliness of grading	2	13.3
Critiques about rubrics/methods of assessment	2	13.3

N=15.

Table 3
Ratings of course in EDUC 774

	Strongly agree	Agree	Neither	Disagree	Strongly disagree
The course content was organized and easy to follow.	18 (58.1%)	12 (38.7%)	1 (3.2%)	0 (0.0%)	0 (0.0%)
The learning objectives were clearly outlined. ^a	20 (66.7%)	9 (30.0%)	0 (0.0%)	1 (3.3%)	0 (0.0%)
The instructional materials supported my learning.	18 (58.1%)	12 (38.7%)	1 (3.2%)	0 (0.0%)	0 (0.0%)
The pace of the course was appropriate.	19 (61.3%)	9 (29.0%)	3 (9.7%)	0 (0.0%)	0 (0.0%)
The technology used in the course supported my learning.	20 (64.5%)	9 (29.0%)	1 (3.2%)	1 (3.2%)	0 (0.0%)
Rate your level of satisfaction with the method of evaluating student learning.	15 (48.4%)	11 (35.5%)	2 (6.5%)	3 (9.7%)	0 (0.0%)
The course was relevant to your professional goals.	20 (64.5%)	9 (29.0%)	2 (6.5%)	0 (0.0%)	0 (0.0%)
Participating with the other online students in the course promoted my learning.	11 (35.5%)	14 (45.2%)	5 (16.1%)	1 (3.2%)	0 (0.0%)
Rate your overall level of satisfaction with the design of the course.	16 (51.6%)	14 (45.2%)	0 (0.0%)	1 (3.2%)	0 (0.0%)

N=31. ^a N=30.

Table 4
Comments about course in EDUC 774

Response	Number	Percentage
Positive comments about the organization of the course	6	46.2
Positive comments about the PDP project	4	30.8
Requests for more variety in the activities	4	30.8
Positive comments about the value of the course (general)	3	23.1
Critiques about discussion boards or PDP project	2	15.4
Critiques about instructor presence	1	7.7

N=13.

Appendix H

ARTIFACT H: UNDERSTANDING BY DESIGN CURRICULUM UNIT

Introduction and problem statement

*A lot of good things are done in the name of professional development.
But so are a lot of rotten things. What educators haven't done is
provide evidence to document the difference between the two.
Evaluation provides the key to making that distinction.*

Guskey, 2000 (p. 94)

This curriculum unit is designed to respond to three challenges. The first is a need to build staff capacity at my organization, the Delaware Education Research and Development Center (DERDC). We will face a 100% turnover among graduate research assistants (GRAs) and recognize the need to bring our new GRAs quickly up to speed on the nature and context of our work. The other two challenges are related to DERDC's work. Most of the programs we evaluate include professional development (PD). While PD *can* be a powerful force for changing teachers' skills and knowledge, improving instruction and, ultimately, student achievement, it often falls short of this goal. There is a clear need to improve PD designs, acting on what has been learned from decades of research. Evaluation can play a role in this improvement, but here we face the third challenge: evaluations of PD also too often fall short of their potential. This unit attempts to address these three issues – the first, directly, and the other two, implicitly – by developing a curriculum for DERDC GRAs, focused on effective PD program evaluations. The essential question for the unit is: *What will it look like when PD is successful?*

Capacity Building at DERDC

DERDC is a center within the University of Delaware College of Education and Human Development that conducts program evaluations, research projects, and policy analyses. DERDC evaluates an array of programs in K-12 education, higher education, and human services. DERDC is a small but busy center, employing two faculty members (one at 40% time) and three professional staff members as well as four GRAs. The staff size has decreased in recent years due to departures and budget efficiencies. Together, this small team is responsible for evaluating approximately 25 programs, the majority of these including PD components.

This month, three GRAs will conclude their terms with DERDC and the fourth will leave in December. We have hired four new GRAs, all doctoral students with some research experience but limited real-world background with schools, PD, and program evaluation. To complete our projects well and on time, it will be essential to orient these GRAs as quickly and thoroughly as possible. To the extent that GRA training processes exist at DERDC, they focus on basic expectations and logistics (e.g., GRAs work 20 hours a week). Beyond this, staff members typically work with GRAs on an as-needed basis, teaching them what they need for each specific task on each specific project. Strengthening this piecemeal approach requires time and planning – hard to find in a busy center, but available to me through this project.

Improving GRA education will benefit our students as well as the center. Feedback from graduating GRAs suggests that DERDC could do a better job developing their skills and giving them opportunities to take on larger roles in our projects, which in turn will strengthen their candidacy for jobs or academic positions. This type of effort is a new venture for DERDC and I built my unit from the ground up. My product is different from a traditional K-12 or higher education curriculum

unit. It is a structure for on-the-job learning about program evaluation and PD. The unit is focused on preparing GRAs to contribute significantly as members of evaluation teams for PD programs. It is highly job-embedded and authentic.

Theoretical Framework and Design Principles

I base my unit on the *Understanding by Design* (UbD) framework (Wiggins and McTighe, 2006). UbD fits my purpose well because it emphasizes understanding, defined as the ability to independently transfer skills and knowledge to an unfamiliar situation. As program evaluators, we do research in the real world, where the unexpected is inevitable. GRAs must learn to think independently and solve problems that come up in the context of evaluating programs. UbD emphasizes complex and authentic tasks. Ultimately, this unit prepares GRAs to demonstrate their understanding by conducting an evaluation of a PD activity within a DERDC project. With its focus on planning backwards from learning goals and objectives, UbD is highly consistent with Tyler (1949, 2013). Indeed, Tyler is also considered one of the pioneers of educational evaluation, with Guskey (2000) building his model on Tyler's ideas.

This unit also integrates adult learning theories (e.g., Merriam, 2001). Our GRAs bring life and professional experience with them. I try to affirm and build on these experiences, respect GRAs as motivated learners, and provide clear applications and clear goals for all tasks. On one level, the relevance of this unit to GRAs is straightforward – it will prepare them to do the job that will fund their education for the next several years. Yet UbD challenges me to deepen this unit so that it is more than basic training but builds understanding about the importance of PD, and of evaluating PD. GRAs may not (yet!) share my passion for these issues. As Wiggins

and McTighe (2006) put it, “the challenge of teaching for understanding is largely the challenge of making the big ideas in the field become big in the mind of the learner” (p. 75). Therefore, my unit features reflection, direct engagement with the literature about PD and PD evaluations, and field experiences – all attempts to move towards understanding.

This unit is built upon a body of literature about PD: why effective PD matters, characteristics of effective PD, and strategies for evaluating and researching PD.

Why effective PD matters

Most educational improvement efforts include educator PD as an essential strategy. As highlighted in our readings this semester, the educational landscape is changing at rapid pace. Technological innovations open new possibilities and demand that teachers change their practices to meet the needs of 21st century learners (Jacobs, 2012). New policies and standards such as the Common Core State Standards (CCSS) or the Next Generation Science Standards (NGSS) require profound shifts in practice, also. Meanwhile, new knowledge about teaching and learning is constantly being produced. The need for teacher learning on every level – classroom, school, system – is acute. Faced with these enormous and rapidly changing demands, those designing educational programs almost inevitably rely on educator PD. As Desimone (2009) puts it, “education reform is often synonymous with teachers’ professional development” (p. 181).

Among others, DERDC is working on evaluations that include PD components for: (a) higher education faculty and secondary science teachers, to support the integration of climate science concepts into their curricula; (b) K-12 teachers, to help them teach argumentation as emphasized by the CCSS; (c) educators at all levels, to

develop assessment items aligned to the Smarter Balanced test; (d) teachers and leaders implementing a cross-district blended learning initiative; (e) educators participating in the statewide World Language Immersion program; (f) midcareer professionals preparing to enter teaching and work in high-need schools. These examples show the range of how PD is used and the prevalence of PD as an improvement strategy. Examples of successful school reform also show the critical role of PD. As Guskey (2000) states, “one constant finding in the research literature is that notable improvements in education almost never take place *in the absence of* professional development” (p. 4).

Characteristics of effective PD

The *No Child Left Behind* legislation (NCLB, 2002) includes a definition of “high quality” PD as activities that (excerpts):

- Improve and increase teacher’s knowledge of the academic subject teachers teach, and enable teachers to become highly qualified;

- Are an integral part of broad schoolwide and districtwide educational improvement plans;

- Are high quality, sustained, intensive and classroom-focused

- Are not one-day or short term workshops or conferences

The law does not provide detailed guidelines for how to design effective PD.

Still, these bullets are a helpful starting point and are generally aligned with the findings of a body of research on what types of PD are most likely to promote changes in teacher skills and knowledge, affect instructional practices, and in turn support increased student achievement.

Although small-scale and case study research about PD is nothing new, since about 2000 there has been an effort to study the effectiveness of PD on student learning through larger, more empirically valid methods, and to look more specifically at what design features of PD make the difference. Recent meta-analyses (Blank and de la Alas, 2009; Yoon, Duncan, Lee, Scarloss & Sharpley, 2007), research syntheses (Desimone, 2009; Borko, 2004) and large studies (Garet, Porter, Desimone, Birman & Yoon, 2001; Penuel, Fishman, Yamaguchi & Gallagher, 2007) have proposed the following features of effective PD:

Content focus – effective PD involves teachers in reinforcing academic concepts, then learning how to convey those concepts to students. So for example, in a math PD, teachers would do math and discuss pedagogical strategies for specific math concepts.

Time/duration – research is clear that teacher learning takes time. For example, of 16 programs identified that demonstrated significant gains in student achievement (Blank & de las Alas, 2009), the average total time was 91 hours and the average duration was 6 months. Some programs had more than 100 hours and lasted 16 months.

Active learning – effective PD engages teachers in a variety of activities that apply concepts to their practice. This could include hands-on activities or lesson planning time during workshops, or less traditional PD experiences such as observing fellow teachers, being observed/coached, participating in school improvement or curriculum development processes. “Sit and get” workshops are the antithesis of effective PD.

Coherence – effective PD attempts to connect with what teachers believe and know already (Desimone, 2009). It also aligns with the existing structures within which they work (e.g., state/national standards, curricula, other reform initiatives).

Collective participation – involving teams of educators from the same system (school, district) has shown to be an effective design. It increases peer support, makes it more likely that changes in instruction will “stick,” and creates momentum for change.

This research basis is also reflected in the national standards for PD, which have also been adopted in Delaware (Learning Forward, 2011). There are standards for learning communities, leadership, resources, data, learning designs, implementation, and outcomes. As those titles suggest, national and state policy emphasize that PD needs to be deeper, more job-embedded, and more collaborative than just a workshop, and we should look more closely at its outcomes.

Strategies for evaluating and researching PD

Despite the perceived importance of PD for promoting changes in instruction and achievement, most research or evaluation on PD does not go far enough. There is a general call for more rigor in research about PD. Of 1300 studies reviewed by Yoon et al (2007), only nine met *What Works Clearinghouse* evidence standards. A chorus of researchers state the need for better designed and more rigorous studies of PD (Borko, 2004; Desimone, 2009). This push is also reflected in the evaluation criteria for federal and state grants and funding. For example, the Delaware Department of Education RFP for the Improving Teacher Quality State Grant Program states that evaluations must “include a formal assessment of the project’s impact on classroom instruction and student achievement” (Delaware Department of Education, 2014).

Current practice in evaluating PD programs often falls far short of these expectations. Most commonly, PD is evaluated by surveying participants to see if they enjoyed the experience. Participant satisfaction is important and necessary but insufficient for understanding the deeper impacts of PD. Reeves (2010) frames the problem starkly: “The central challenge for educational systems around the world is the substitution of effectiveness for popularity” (p. 2). To make this move, we can consider the five levels of evaluation proposed by Guskey (2000). There are others in the field of evaluating PD (e.g., Killion, 2007) but Guskey’s model is the most influential.

According to Guskey, evaluation of PD should attend to five levels of effects: participants’ reactions, participants’ learning, organization support and change, participants’ use of new knowledge and skills, and student learning outcomes (see Appendix 2). For Guskey, each level is a pre-requisite for the others yet success at one level does not necessarily mean success at the next. For example, a participant can love a PD experience and learn a lot in it (Levels 1 and 2), but if his/her school or district context fails to support the new learning (Level 3), s/he may never demonstrate changes in teaching practice (Level 4) and student outcomes may not improve (Level 5). This could happen if structures such as the school schedule, teaching/PLC arrangements, or curriculum requirements compromise the implementation of PD, or if school leaders or fellow teachers did not understand or support it. (This is an issue of coherence). To understand the intricacies of PD transfer, it is necessary to evaluate all five levels. As the levels increase, evaluation becomes more complicated and time-consuming and yet the information gathered is more valuable. Guskey argues that even

organizers of small PD programs can inquire into all five levels, and external evaluators certainly should.

At DERDC, we generally follow the Guskey model but our ability to do so depends on the project budget and timeline. In multi-year projects, we have the opportunity to examine not only the PD activities themselves and how participants respond to them and learn from them, but also what changes, if any, happen back at school in terms of organizational culture, teaching practices, and student achievement. Sometimes, however, project timelines or budgets preclude such a thorough evaluation. Ironically, despite the evaluation requirements cited above, DDOE requires all Improving Teacher Quality funds to be spent by July 31, making it impossible to examine transfer of PD to the classroom. Looking forward, DERDC hopes to increase and institutionalize the use of the Guskey framework. Increasing understanding of this framework among our GRAs is an important step for capacity building.

Strategy

Wiggins and McTighe propose six “doorways to design” for a UbD unit. I walked through the second door: “begin by considering desired real-world applications.” In this case, the application is conducting evaluations of PD programs and the goal is to have GRAs ready to participate as members of evaluation teams as soon as possible. Still, I also recognize that teaching for understanding takes longer. I could have chosen a more directive approach, teaching GRAs to “plug and play” without asking too many questions about what they were doing, or why. This may have been faster, but it would not have built GRA understanding, and thus DERDC capacity, enough. It would also have failed to treat GRAs as adult learners with relevant experiences and motivations of their own. I hope that this investment in GRA

education at the beginning of their DERDC employment will pay dividends over the four years they are with us.

Since this unit is on the job training, opportunities for real-world application abound. Initially, I had planned for this unit to involve GRAs in a sequence of tasks with DERDC evaluations, building up to a performance task in which GRAs conducted all the steps of a simple evaluation. Using the UbD framework helped me recognize that I needed to add a performance task, and associated learning activities, that looked beyond our work at DERDC. At DERDC, we do our work for clients, for a price, with fixed schedules and evaluation plans. This means that GRAs would have limited opportunities on our actual projects to do some of the steps (e.g., design evaluation plans) that would be necessary for building understanding. Since making mistakes in an evaluation can be detrimental to clients, DERDC is fairly conservative in the tasks we assign to GRAs. This is understandable, and yet if we do not involve them more deeply in the work, our capacity will not increase. Therefore, in addition to a job-embedded performance task, I also created one that did *not* involve actual projects, for which I could use a more constructivist approach and let our GRAs think more creatively and make more interesting mistakes. I worked backwards from these two performance tasks to develop other learning experiences.

This unit teaches GRAs to evaluate PD using the Guskey model, and to think about the features of effective PD according to the research discussed above, which is reflected in state and national PD standards. As far as I know, there are no standards for GRAs at UD. While standards for program *evaluations* exist and are widely used (Yarbrough et al, 2011), there is a more recent movement to define essential competencies for program *evaluators*, that is the individuals doing the work (Stevahn

et al., 2005). There are six categories of competencies: professional practice, systematic inquiry, situational awareness, project management, reflective practice, and interpersonal competence (see Appendix 1). While no unit for GRAs could cover *all* 61 competencies, I tried to identify the most salient ones for brand new program evaluators, and to align my unit with them.

Product

This is a brand-new unit, not a revision. The UbD template for my unit is found in Appendix 3. The essential question for my unit is: *What will it look like when PD is successful?* The enduring understandings are that PD matters, PD is not magic, and evaluation can help. These are all derived directly from the research discussed above. This research basis, along with the Guskey model and sample evaluation reports, comprise the content knowledge of this unit. Skills taught include basic program evaluation skills such as creating a logic model, matching evaluation questions to data sources, collecting, analyzing, and interpreting data. These all fall under the essential competency of systematic inquiry.

The culmination of the unit, Performance Task 2 is to evaluate a PD activity within a DERDC-evaluated project. DERDC staff will provide oversight, but GRAs take the lead. The activity will be chosen based on project schedules, GRA schedules, and GRA readiness. Most likely it will be a one-or two-day PD workshop. Specific components of this task include revising data collection instruments, collecting data (completing observation protocol, administering surveys) at the event, analyzing this data, and interpreting it in a brief report. In order to reflect on this process and connect it to prior learning, GRAs will engage in discussion with DERDC staff. They will put the evaluation of the PD activity in the context of the broader project, considering

what Guskey level(s) are considered through the evaluation they conducted, and what level(s) will be considered through future evaluation activities for the project.

Completed earlier, Performance Task 1 asks GRAs to be creative and imagine they are the Director of Instruction in a district, preparing a grant proposal and evaluation plan for a new PD initiative whose overall goal is to improve student achievement in either non-fiction writing or engineering design. GRAs draw a logic model for their program and answer questions: who participates? What activities does the PD program include? What are the expectations for teacher participants? They then describe how their program is aligned with the research on effective PD including one new article that they locate. Finally, they outline their evaluation plan including evaluation questions, data sources, data collection method/instruments, and timelines. This does not have to be a formal write up; even bullets would be sufficient. The point is for GRAs to show that they understand what an exemplary research-based program and evaluation would look like.

The rest of the unit builds up to these performance tasks and includes a number of tasks (noted with letters in the UbD framework). It begins with an activity to activate GRA's prior knowledge and to help GRAs build community. GRAs journal on the *most* and *least* effective PD experiences they have participated in, then share responses with a colleague, report out, and synthesize across responses. Then, GRAs are given a structured set of questions about PD design (Guskey, 2000) and complete these for both their most and the least effective experience. The debrief for this activity will be metacognitive. How does changing the "instrument" (i.e. open-ended journal vs. structured checklist) affect the type of information received? When might you want

to use each method? This is a simple way to begin GRAs thinking about evaluation design.

GRAs will read the literature on effective PD. This is an opportunity for differentiation. GRAs with more interest in or prior experience with PD or with scholarly research may read journal articles (e.g., Desimone, 2009). Others may, at least to start, read practitioner articles (e.g., Desimone, 2011). They will also read Guskey (2000). To begin to see how this literature translates into practice, GRAs will read an example evaluation report of a PD program that uses the Guskey model (Wolanin & Wade, 2013 or Munoz, Guskey & Aberli, 2008). They will identify all five levels, noting what types of evaluation questions are asked at each one and what data collection methods and data sources are employed. This activity responds to GRA feedback that they would like more models of evaluation reports.

The unit also includes a set of activities in the field where GRAs gain direct experience with PD. Most of these involve attending activities within projects that DERDC is evaluating. The GRA will take on gradually increasing responsibilities. First s/he will shadow a DERDC colleague through the process of attending and collecting data at the session, analyzing and interpreting the data, and writing a brief report. Then s/he will have the opportunity to practice each of these steps. Over time, the level of detail and complexity will grow. For example, the GRA could first take field notes, then complete an observation instrument and discuss/compare scores with a colleague. The GRA would first simply administer surveys, but over time also revise the instruments to reflect the specifics of the PD activity, and conduct straightforward quantitative analyses of ratings items and qualitative coding of open-ended responses. GRAs are often fairly adept at quantitative analyses. They struggle more with

interpreting these analyses since they often lack a background understanding about education and PD. Hopefully this unit will improve that issue. Qualitative analysis is often new to them, and analyzing open-ended survey responses is a good first step.

The unit also includes an opportunity (Task F) for GRAs to independently attend and report back on a PD event outside of DERDC. This could be a workshop or training for graduate students (e.g., one offered through UD) or another event they identify. If needed, I could help GRAs locate opportunities. The goal in sending GRAs out alone is to see whether they can transfer and apply their growing understanding about PD to a new event, and whether they can make independent judgments about it. Of course, it will take a while before GRAs are ready to conduct actual DERDC evaluations autonomously. In our current group, only GRAs in their third and fourth year with us are at this level of responsibility. However, if the goal of my unit is to develop understanding, GRAs need chances to make sense of what they see on their own.

In some important ways, my unit is unlike a unit that would be taught in K-12 or higher education. It does not include traditional lesson plans but rather is a set of tasks that build on each other and through which the GRA would move. I have sequenced them, except for Task G (Scaffolded DERDC Evaluations) which would be spread throughout the unit. The order of events may depend on actual project schedules. However, unlike the “playlists” we learned about in Jacobs (2012) and the BRINC project, this cannot be a totally student determined, self-paced unit. For example, it is necessary to know something about PD before learning the Guskey model. GRAs have to demonstrate their professionalism and ability to collect data at PD before they can go out on their own or be responsible for an evaluation,

Since this type of GRA education is so new to DERDC, it is difficult for me to predict how long this unit will take. GRAs are likely to move through it at different paces. The most collaborative activities come at the start, when GRAs discuss prior PD experiences together and work in pairs to review research on PD. As they begin to work on real DERDC projects, they are likely to split up. Because we have many projects and a small staff, it is unusual for more than one GRA to work on the same project. Most of the “teaching” of this unit will take place through mentoring relationships between GRAs and the DERDC staff member working on the same project. However, in order to learn from each other’s experiences and build relationships, I hope that we can also have regular GRA meetings to discuss how the work is going. Even if GRAs are tackling different tasks within this unit, they should still have plenty to talk about.

Another difference between this unit and a school unit is that we do not grade GRAs. However, my unit does include assessment rubrics for Performance Tasks 1 and 2 (see Appendix 4), breaking out the key components of each task and describing levels of proficiency. In addition, GRA competence with basic evaluation skills (data collection, analysis and interpretation) will be continuously evaluated during the unit. Taken together, these rubrics also could provide the basis for developing a more robust performance evaluation system for GRAs. That is beyond my scope here but has been identified as a need by staff and GRAs alike.

This project has several limitations. First, it does not include dates, for the reasons described above. After going through this unit once, we may develop a better sense of the timing. Second, I do not have much information about our incoming GRAs beyond what I learned in brief interviews and on their resumes. I planned this

unit based on what would have been helpful for our current GRAs, according to their feedback and my impressions. Yet the new students may have very different abilities or interests, which would require adjusting the unit accordingly. Third, there is still a need to develop more specific plans for teaching quantitative and qualitative data analysis and evaluation report writing. Within my unit, this would fall during Task G. We are having internal conversations about the need to develop nuts-and-bolts workshops/mini-lessons on topics such as: cleaning and entering data, using SPSS to run descriptive statistics, when and how to use bivariate statistics, coding open-ended responses, and presenting coded data. We also see the need to model and teach how we write reports at DERDC, including APA style. Such detailed plans are important for my unit and would be the next stage of development.

Reflection

I began planning with a sense of urgency about our imminent staff turnover at DERDC. As I developed my unit, I realized how time consuming and demanding capacity building can be. As described above, my choice to use UbD and focus on transfer skills meant that I had to take a more in-depth approach. I hope that this investment in GRAs will pay off over the several years they are with DERDC, but at the outset the unit may not feel like it is moving fast enough.

Conversations with my EDUC 897 classmates also made me realize the need to persuade different stakeholders about this approach to educating GRAs. These stakeholders include my colleagues, my supervisor, clients, and the GRAs themselves. Pressured by timelines and deliverables, colleagues may be reluctant to invest time into working with GRAs. Delegating and capacity building sometimes seems more time consuming than doing the task oneself. This is also a trap I need to watch out for

myself. Sometimes DERDC will want to use GRAs for time-sensitive administrative tasks, cutting into the time available for this unit. It will also be important to gain my director's support for this approach. For example, we would be asking GRAs to attend a non-DERDC PD activity during work hours. This may raise complications for accounting/paying for their time. In general, staff members whose projects we evaluate are open to having GRAs helping, but we may need to explain our approach because most likely GRAs will be more visible than they have in the past.

Some of my classmates thought this unit was too ambitious and asked too much of GRAs. I disagreed, and in thinking through my reaction, I realized how strongly I believe in developing GRAs. Some of this is just my style. I was also a challenging teacher and non-profit director, and I believe that students will rise to whatever standards you set for them. But it also comes from personal experience from being a GRA at DASL. As a first-year masters student, I presented to 100+ people at the statewide Policy and Practice Institute and gave a one-on-one briefing to the Secretary of Education. I had these opportunities because my supervisors believed in learning by doing and routinely gave GRAs complex, authentic, and visible tasks. At times I felt over my head, and I never had the benefit of a training "curriculum" like this, but overall, I was grateful. When I in turn supervised GRAs at DASL, I tried to push them forward as well. The culture and the work at DERDC are different, and in general we are more risk averse about letting GRAs take responsibility and possibly make mistakes. This is understandable given the more high-stakes nature of our work. I think my unit strikes an appropriate balance between my GRA experience at DASL and current practice at DERDC. At the same time, because it is new for the organization, I may experience push back from GRAs who just want to work their 20

hours a week and not tackle larger, real-world evaluation responsibilities. I am hoping, however, they will find the unit engaging and worthwhile, and be eager to build the understanding they will need to carry more responsibilities during their GRA experience.

References

- Blank, R.K. & de las Alas, N. (2010). Effects of teacher professional development on gains in student achievement: how meta analysis provides scientific evidence useful to education leaders. Society for Research on Educational Effectiveness.
- Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educational Researcher*, 33 (8), 3 – 15.
- Delaware Department of Education (January 3, 2014). Highly qualified teachers and improving teacher quality state grant program [Request for proposal].
- Desimone, L. (2011). A primer on effective professional development. *Phi Delta Kappan*, 92 (6), 668 – 671.
- Desimone, L. (2009). Improving impact studies of teachers' professional development: Towards better conceptualizations and measures. *Educational Researcher*, 38 (3), 181 – 199.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F. & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38 (4), 915 – 945.
- Guskey, T. R. (2000). *Evaluating professional development*. Newbury Park, CA: Corwin Press.
- Jacobs, H.H. (2012). *Curriculum 21: Essential education for a changing world*. Alexandria, VA: ASCD.
- Killion, J. (2007). *Assessing Impact: Evaluating Professional development*, 2nd edition. Newbury Park, CA: Corwin Press.
- Learning Forward (2011). Standards for professional development. Retrieved April 21, 2014 from <http://learningforward.org/standards>

- Merriam, S.B. (2001). Andragogy and self-directed learning: Pillars of adult learning theory. *New Directions for Adult and Continuing Education*, 89, 3-14.
- Munoz, M.A., Guskey, T.R., & Aberli, J. (2008). *Struggling Readers in Urban High Schools: Evaluating the Impact of Professional Development in Literacy*. Louisville, KY: Jefferson County Public Schools. Downloaded May 19, 2014 from <http://goo.gl/iaM0Hk>.
- No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).
- Penuel, W.R., Fishman, B.J., Yamaguchi, R. & Gallagher, L.P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, 44 (4), 921 – 958.
- Reeves, D. (2010). *Transforming Professional Development into Student Results*. Alexandria, VA: ASCD.
- Stevahn, L., King, J.A., Ghore, G., & Minnema, J. (2005). Establishing essential competencies for program evaluators. *American Journal of Evaluation*, 26 (1), 43 – 59.
- Tyler, R. (1949, 2013). *Basic principles of curriculum and instruction*. Chicago, IL: University of Chicago Press.
- Wiggins, G. & McTighe, J. (2006). *Understanding by design* (2nd Ed.) Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- Wolanin, N.L. & Wade, J.H. (2013). *Evaluation of the Science, Technology, and Engineering*
- Leadership Program, Year Two*. Rockville, MD: Montgomery County Public Schools. Downloaded May 18, 2014 from <http://goo.gl/l2Xyuc>.
- Yarbrough, D.B., Shulha, L.M., Hopson, R.K., & Caruthers, F.A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users*. Thousand Oaks, CA: SAGE.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007).
Reviewing the evidence on how teacher professional development affects
student achievement (Issues & Answers Report, REL 2007–No. 033).
Washington, DC: U.S. Department of Education, Institute of Education
Sciences, National Center for Education Evaluation and Regional Assistance,
Regional Educational Laboratory Southwest. Retrieved from
<http://ies.ed.gov/ncee/edlabs>

Appendix 1
Essential competencies for program evaluators
(Stevahn, King, Ghore & Minnema, 2005)

- 1.0 Professional practice
 - 1.1 Applies professional evaluation standards
 - 1.2 Acts ethically and strives for integrity and honesty conducting evaluations
 - 1.3 Conveys personal evaluation approaches and skills to potential clients
 - 1.4 Respects clients, respondents, program participants, and other stakeholders
 - 1.5 Considers the general and public welfare in evaluation practice
 - 1.6 Contributes to the knowledge base of evaluation
- 2.0 Systematic inquiry
 - 2.1 Understands the knowledge base of evaluation
 - 2.2 Knowledgeable about quantitative methods
 - 2.3 Knowledgeable about qualitative methods
 - 2.4 Knowledgeable about mixed methods
 - 2.5 Conducts literature reviews
 - 2.6 Specifies program theory
 - 2.7 Frames evaluation questions
 - 2.8 Develops evaluation designs
 - 2.9 Identifies data sources
 - 2.10 Collects data
 - 2.11 Assesses validity of data
 - 2.12 Assesses reliability of data
 - 2.13 Analyzes data
 - 2.14 Interprets data
 - 2.15 Makes judgments
 - 2.16 Develops recommendations
 - 2.17 Provides rationales for decisions throughout the evaluation
 - 2.18 Reports evaluation procedures and results
 - 2.19 Notes strengths and limitations of the evaluation
 - 2.20 Conducts meta-evaluation
- 3.0 Situational analysis
 - 3.1 Describes the program
 - 3.2 Determines program evaluability
 - 3.3 Identifies the interests of relevant stakeholders
 - 3.4 Serves the information needs of intended users
 - 3.5 Addresses conflicts
 - 3.6 Examines the organizational context of the evaluation
 - 3.7 Analyzes the potential considerations relevant to the evaluation
 - 3.8 Attends to issues of evaluation use
 - 3.9 Attends to issues of organizational change
 - 3.10 Respects the uniqueness of the evaluation site and client
 - 3.11 Remains open to input from others
 - 3.12 Modifies the study as needed
- 4.0 Project management

- 4.1 Responds to requests for proposals
- 4.2 Negotiates with clients before the evaluation begins
- 4.3 Writes formal agreements
- 4.4 Communicates with clients throughout the evaluation process
- 4.5 Budgets an evaluation
- 4.6 Justifies cost given information needs
- 4.7 Identifies needed resources for evaluation, such as information, expertise, personnel, instruments
- 4.8 Uses appropriate technology
- 4.9 Supervises others involved in conducting the evaluation
- 4.10 Trains others involved in conducting the evaluation
- 4.11 Conducts the evaluation in a nondisruptive manner
- 4.12 Presents work in a timely manner
- 5.0 Reflective practice
 - 5.1 Aware of self as an evaluator (knowledge, skills, dispositions)
 - 5.2 Reflects on personal evaluation practice (competencies and areas for growth)
 - 5.3 Pursues professional development in evaluation
 - 5.4 Pursues professional development in relevant content areas
 - 5.5 Builds professional relationships to enhance evaluation practice
- 6.0 Interpersonal competence
 - 6.1 Uses written communication skills
 - 6.2 Uses verbal/listening skills
 - 6.3 Uses negotiation skills
 - 6.4 Uses conflict resolution skills
 - 6.5 Facilitates constructive interpersonal interaction
 - 6.6 Demonstrates cross-cultural competence

Appendix 2
Five critical levels of professional development evaluation
(Guskey, 2000)

<i>Evaluation level</i>	<i>What questions are addressed?</i>	<i>How will information be gathered?</i>	<i>What is measured or assessed?</i>	<i>How will information be used?</i>
1 – Participants' reactions	Did they like it? Was their time well spent? Did the material make sense? Will it be useful? Was the leader knowledgeable & helpful? Were the refreshments fresh & tasty? Was the room the right temperature? Were the chairs comfortable?	Questionnaires administered at the end of the session Focus groups Interviews Personal learning logs	Initial satisfaction with the experience	To improve program design and delivery
2 – Participants' learning	Did participants acquire the intended knowledge and skills?	Paper and pencil instruments Simulations and demonstrations Participant reflections Participant portfolios Case study analyses	New knowledge and skills of participants	To improve program content, format, and organization
3 – Organization support and change	What was the impact on the organization? Did it affect organizational climate and procedures? Was implementation advocated, facilitated, and supported?	District and school records Minutes from follow-up meetings Questionnaires Focus groups Structured interviews with participants and school or district administrators	The organization's advocacy, support, accommodation, facilitation and recognition	To document and improve organizational support To inform future change efforts

	Was the support public and overt? Were problems addressed quickly and efficiently? Were sufficient resources made available? Were successes recognized & shared?	Participant portfolios		
4 – Participants' use of new knowledge and skills	Did participants effectively apply the new knowledge and skills?	Questionnaires Structured interviews with participants and their supervisors Participant reflections Participant portfolios Direct observations Video or audiotapes	Degree and quality of implementation	To document and improve the implementation of program content
5 – Student learning outcomes	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?	Student records School records Questionnaires Structured interviews with students, parents, teachers and/or administrators	Student learning outcomes -Cognitive -Affective -Psychomotor	To focus and improve all aspects of program design, implementation, and follow up To demonstrate the overall impact of professional development

Appendix 3
Understanding by Design Template

Stage 1 Desired Results		
ESTABLISHED GOALS/ STANDARDS Essential competencies for program evaluators in: - Professional practice - Systematic inquiry - Situational analysis - Project management - Reflective practice - Interpersonal competence	Transfer	
	<i>Delaware Education Research and Development Center (DERDC) Graduate Research Assistants (GRAs) will be able to independently use their learning in new situations to...</i>	
	Evaluate professional development (PD) programs in ways that demonstrate an understanding of the elements of effective PD, the five levels of evaluating PD, and the basics of program evaluation.	
	Meaning	
	UNDERSTANDINGS <i>Students will understand that...</i> PD matters. All educational improvement stories include PD as a key strategy. We care about PD because it can help us build better schools where teachers teach better and students learn more. PD is not magic. Although there is a research basis for how PD leads to improvements in teaching and learning, it is often hard to translate to practice. Evaluation plays a role in bridging this gap. Evaluating PD can help. Most DERDC projects include PD. Our purpose is to ask questions, collect and analyze data, and form recommendations to improve the PD program so it supports educator and student learning.	ESSENTIAL QUESTIONS: <i>Students will explore & address these recurring questions:</i> What will it look like when PD succeeds?

<i>Acquisition</i>		
	<i>Students will know...</i>	<i>Students will be skilled at...</i>
	State and national standards for professional development	Identifying/searching for relevant literature
	Research consensus on characteristics of effective PD	Reading and synthesizing literature about PD
	Basic steps in an educational program evaluation	Understanding the components of a program evaluation (via DERDC and other exemplars)
	Guskey's (2000) five-level model for evaluating PD	Creating a logic model for a sample PD program
	Simple quantitative and qualitative data analysis methods	Matching evaluation questions to possible data sources to answer those questions
	What an evaluation report should look like and what information it should include	Collecting data from a PD program through observation, document review, surveys, or interviews.
		Analyzing data from a PD program through basic qualitative and quantitative methods
Stage 2 - Evidence		
Alignment Coding	Evaluative Criteria	Assessment Evidence
		PERFORMANCE TASK(S): (see more details on both below) Task 1: Fantasy program design and evaluation plan – GRAs will design a hypothetical PD program that illustrates the research basis on features of effective PD. They will also design an evaluation plan for this program that exemplifies all levels of the Guskey framework. Resources are not a constraint. Task 2: DERDC evaluation piece – GRAs will take the lead in evaluating a PD activity within one of DERDC's projects, within the constraints of the evaluation plan/budget. Tasks include instrument revision, data collection via at least observations and surveys, data analysis, and interpretation. The final step is a reflection on whether the evaluation demonstrates Guskey's framework.

	OTHER EVIDENCE:
	• Synthesis of PD experience • Summary of research on PD • Identification of an appropriate PD research article • Identification of Guskey levels in an evaluation report • Detailed field notes aligned with concepts discussed in the unit (effective PD design, Guskey)
Stage 3 – Learning Plan	
Coding	<i>Summary of Key Learning Events and Instruction (including pre- and formative assessments)</i>
	A: Reflections on PD experiences Competencies targeted: aware of self as an evaluator; uses verbal/listening skills GRAs will be able to: • Describe from experience whether a PD/training experience was effective or not and <i>why</i>; • Listen to and summarize a partner's experiences; • Synthesize and summarize any themes across numerous responses; • See that the way you ask a question affects the information you get in response. Introduce the essential questions of this unit and the goal of preparing GRAs to develop program evaluation skills and understanding of PD contexts that they can transfer to actual project work. Next, ask them to reflect on recent professional training or PD experiences they have had – the <i>most</i> and the <i>least</i> effective. Journal in open-ended format. Ask GRAs to share with a partner (facilitator chooses whether they share the 'most' or the 'least' effective first) and report out on key points from what <i>their partner</i> told them. Discuss: were there common themes across the experiences? What were the differences between the "most" and the "least" effective experiences? Next, ask GRAs to either complete the questionnaire on Guskey (2000) p. 70. This asks a set of structured questions. Again debrief. Consider whether changing the 'instrument' affects the kind of information that we get on these questions. What could you do with the questionnaire data that you could not do with the open-ended debriefs? "Pull the curtain back" to reveal that we just touched on several issues that this unit will tackle – what makes PD effective

	(or not) and what are different ways to gather information on how effective it is, and why. Finally, ask GRAs to complete basic information describing their prior experience in PD (as participant, facilitator, researcher/evaluator or other role) and their one biggest question about PD/thing they want to learn about PD Formative assessment: note the level of engagement/participation in the discussion B. Laying the foundation of research on PD Competencies targeted: conducts literature reviews, develops logic models GRAs will be able to: • Read summaries of PD research and discern key findings of each • Develop a generic logic model for a PD program • Locate relevant, recent literature through database search • Understand the rationale for state and national PD standards Read Introduction and Chapters 1, 2 & 3 of Guskey (2000) and work by Desimone (2009 or 2011), Borko (2004) & Blank & de las Alas (2010). Differentiation: based on GRA time, interest, and degree of prior experience with education research, they may be guided towards full research articles or shorter practitioner articles. Restate essential question: what will it look like when PD succeeds? Discussion: How have these authors contributed to our understanding of this question? What do we know now? At the most simple, how is PD for educators supposed to work? Draw a simple flow chart. Optional extension activity: for GRAs who are ready for the challenge or already have the experience, talk about developing a logic model for a sample PD program (for example, the ones described by Blank & de las Alas, 2010) Next, ask GRAs to conduct a database search. (If necessary, provide them with access to a UD Library webinar on effective
--	---

	search strategies). Ask them to find one article discussing research/evaluation of a PD program and its link to student achievement. They will use this article in performance task 1. C. PD Observation By this point, if it has not already occurred, ensure that GRAs are able to attend an actual PD event. Role will be simply to observe. After observing PD session, have open-ended discussion and debrief with GRA: what did you see that seemed to generally fit or not fit with what you have learned so far about research-based PD practices? GRA also can “shadow” DERDC staff member through the process of analyzing and interpreting data and writing the report. D. The Guskey evaluation framework Competencies targeted: understands the knowledge base about evaluation GRAs will be able to: • Know and apply Guskey’s five-level PD evaluation method • Align evaluation questions, data collection instruments, and data sources in sample evaluation report to Guskey’s five levels • Transfer their growing understanding of PD and program evaluation to a fictional program design and evaluation plan If needed, briefly recap differences between evaluation and research. GRAs read Guskey (2000) chapters 4 – 8. GRAs break into two groups and read one of these two evaluation reports. Can divide by academic background: • <i>Evaluation of the Science, Technology and Engineering Leadership Program, Year 2</i> (Wolanin & Wade, 2013) • <i>Struggling Readers in High School: Evaluating the Impact of Professional Development in Literacy</i> (Munoz, Guskey & Aberli, 2008) Use colored pencils/highlighters to identify evaluation questions, data sources, and data collection instruments that correspond to each level.
--	---

	Questions: Did the evaluation look at all five Guskey levels? Guskey argues that as the levels go up, conducting the evaluation becomes more challenging but the results are more meaningful. Do you see evidence of this in the reports? E. Performance Task 1 Fantasy program and evaluation design (Caveat: this is not the real world....!) Competencies targeted: Frames evaluation questions; develops evaluation designs; develops logic models You are the director of instruction in a school district, preparing a grant proposal for a new initiative. The goal is to improve student achievement in a particular subject area. You may choose from these subjects, newly emphasized in the NGSS or CCSS: engineering design or informative/argumentative writing. The centerpiece of this initiative is PD for teachers across the district. You are well versed in the research about effective PD and you have plenty of resources. You are also working with an external evaluator to figure out how to evaluate the initiative and its impact on teachers' knowledge and skills, their instructional practices, and student achievement. Your evaluator knows the Guskey framework. Draw a logic model for your program. Then provide more details about the design of your PD program, answering questions such as: • Who participates in this PD program? • What activities does the PD program include? • What is the schedule? • What are the expectations for teachers participating in the program? Include a brief description of how your design illustrates the research consensus on features of effective PD. Include references to articles we read together as well as at least one other article that you locate. Next, outline your evaluation plan including evaluation questions, data sources, data collection method/instrument, and timeline.
--	---

	Do not worry about trying to write an official “grant proposal” or “evaluation plan.” This can even be bullets or tables. The important thing is to show that you understand what an exemplary research-based program and a Guskey-based evaluation would look like. F. Independent PD review Competencies targeted: describes the program; uses verbal/listening skills GRAs will be able to: • Conduct themselves professionally and independently in observing or participating in a PD experience • Take detailed field notes • Align field notes with the design features of effective PD To “return to the real world” and continue to diversify their PD experiences, GRAs are asked to identify an opportunity to observe a PD session for educators in some context. This should not be part of a DERDC project. GRAs attend the PD and take field notes; then afterwards describe the program (verbal debrief is ok). Specifically, they should be identifying to what extent the event demonstrates aspects of highly effective PD. G. Scaffolded DERDC evaluations Competencies targeted: knowledgeable about quantitative methods, knowledgeable about qualitative methods, collects data, analyzes data, reports evaluation procedures and results GRAs will be able to: • Assume gradually increasing responsibility for evaluating a DERDC project PD activity This activity is ongoing throughout the unit and will be coordinated with appropriate project activities. GRAs will attend PD activities with DERDC staff members. Their role will build gradually (e.g., first they just observe; next they observe, take notes, and complete observation protocol; next they assist with survey administration, etc.). It is anticipated that GRAs will attend at least three PD activities over the course of the unit, with increasing responsibilities in each one. Individual timelines and plans will be developed for each GRA based on background, interests, and schedule availability. The DERDC staff member working with each GRA will determine when s/he is ready for Performance Task 2.
--	--

H. Performance Task 2: Actual DERDC Evaluation

This task is real life, not a scenario. You are a DERDC GRA conducting an evaluation of a PD activity. The PD activity will be chosen based on project schedules and your availability, but it will probably be a one or two-day workshop. Examples may include: MADE CLEAR teacher PD (fall 2014) or higher education workshop (August 2014); Sustaining Places workshop (fall 2014); DT₃P candidate workshop (fall/winter 2014-5).

Your tasks include:

- Revise data collection instruments (almost certainly a survey, possibly also interview protocols) based on information about the specific activity and its goals and objectives.
- Collect data at the activity. Depending on the evaluation plan, this could include: complete observation protocol, take field notes, administer surveys, interview facilitators or participants.
- Analyze data from the activity. This requires quantitative analysis of closed-ended survey items and qualitative analysis of open-ended survey items and any interview data.
- Interpret data and participate in discussions about the data with DERDC staff (who probably know the program context better). Compile an evaluation report.
- Apply the Guskey framework and prepare to discuss with DERDC staff members: what levels are we evaluating here? Review the broader evaluation plan for the project. Will there be an evaluation of deeper levels? If so, how and when? If not, why not? And what might a more thorough evaluation look like? What questions would it answer?

Appendix 4
Rubric - Performance Task 1

	Mastering	Progressing	Beginning
Design of PD	Proposed program includes all components of effective PD and demonstrates careful thought and understanding. Logic model is clear and easy to understand.	Program includes most components of effective PD <i>and/or</i> components are inserted without careful connections to the purpose. Logic model is adequate.	Proposed program includes some of the components of effective PD <i>and/or</i> includes design elements that are <i>not</i> research supported. Logic model may contain errors.
Research connections	GRA has identified an appropriate, credible article, explained its relevance for the project, and connected it clearly to other works.	GRA has identified an article and explained its relevance. Links to prior research may be unclear or cursory and/or article may be less credible	Some aspect of the task is incomplete (e.g., the article does not meet criteria, previous research is not mentioned, links to PD program may not be made)
Evaluation plan	Evaluation plan includes all five levels of Guskey framework with appropriate questions posed at each level. Data collection methods and data sources fit the questions.	Evaluation plan includes all five Guskey levels. Some of the questions may be misplaced and/or evaluation questions, data collection methods and data sources do not align.	Evaluation plan does not include all five Guskey levels or includes misunderstandings about the levels.
Reflection	Reflection indicates that GRA can “reality check” the PD proposal and evaluation plan and can also integrate his/her own PD experiences.	Reflection includes some recognition of constraints to PD programs or evaluations, and some connection to personal experience.	Reflection is cursory or non-existent.

	Mastering	Progressing	Beginning
Data collection	Data collection instruments reflect specific PD agenda. Observation data are detailed, accurate and show understanding of PD.	Data collection instruments are revised. Observation data are collected with care.	Data collection instruments not revised or not appropriate and/or observation data insufficient or off base.
Data analysis	Quantitative and qualitative analyses are accurate; data are displayed appropriately.	Data analyses are generally accurate. There may be an error or two.	Data analyses include several errors or inaccuracies.
Data interpretation	Discussion of data shows understanding of PD. Meaningful recommendations are provided.	Discussion is related to the data and appropriate for the purpose. Recommendations are given but may be cursory.	Discussion not anchored in the data or simply restates the data. No recommendations are given.
Report writing	Report is well organized, with clear, concise prose. Report follows APA and DERDC style.	Report generally well organized and correctly written, though some revisions may be necessary.	Report is hard to follow. Sentences and paragraphs are not clear. Report does not follow APA and DERDC style.
Reflection	Reflection shows a clear understanding of Guskey and of the broader PD project for which an activity was evaluated.	Reflection shows some efforts to contextualize the PD activity and apply the Guskey framework.	Reflection contains misunderstandings about Guskey or limited evidence of thought or application.
Professional practice	GRA is consistently on time, prepared, appropriately dressed when attending project activities. S/he is respectful & unobtrusive.	GRA is usually on time, prepared, appropriately dressed, respectful & unobtrusive.	There may be issues in punctuality, preparation, appropriate dress or behavior.

Appendix I

ARTIFACT I: ONLINE PROFESSIONAL DEVELOPMENT TOOLKIT

This artifact is a website. It can be accessed at the following URL:

<https://sites.google.com/a/udel.edu/online-pd-toolkit/?pli=1>

Appendix J

ARTIFACT J: ANALYSIS OF STUDENT PROFESSIONAL DEVELOPMENT PLANS FROM EDUC 774

Introduction

This analysis examines professional development plans (PDPs) created as final projects for a master's course on Designing Professional Development. These plans provide insight into how a group of aspiring teacher leaders envision professional development (PD) that meets their schools' priorities and their own professional learning needs, choices and values. When analyzed with a structured framework built around research-based components of effective PD, the plans indicated alignment with a number of characteristics but not all. Further, the plans varied substantially in the level of depth and detail with which they integrated research-based features.

Method

Plans for this study were submitted to the University of Delaware's Institutional Review Board and granted exempt status. All 58 students who had been enrolled in EDUC 774 in spring 2016 were informed of the study in November 2016, using both email and Canvas messages. The class included mostly masters' students in the teacher leadership program with a few doctoral candidates as well. They came from varied educational settings including colleges/universities and K-12 schools in Delaware or the surrounding states. Using a passive consent process, one student out of 58 (1.7%) opted out of participation. I removed all identifying information and

numbered the PDPs. For alignment with the rest of my ELP, I decided to focus this analysis on PDPs written for elementary, middle or high schools in Delaware; 33 plans met this description. Of these, three (9.1%) planned PD for multiple grade levels and the others were balanced between elementary (14, 42.4%) or secondary (16, 48.4%) education. Of the secondary plans, half (8) focused on middle school and half (8) on high school

The framework used to analyze the PDPs was developed from three overlapping sources: (1) the Learning Forward Standards for Professional Learning (Learning Forward, 2011); (2) a research-based framework of features of effective PD (Desimone, 2009); and (3) the course syllabus, including a template provided to students to scaffold the design of their PDPs. Ultimately the framework included ten characteristics, defined in Table 1.

Table J1 Features of professional development

Feature	Source	Definition
Content focus	Desimone (2009), syllabus/template	The PD focuses on content knowledge and how students learn it (i.e., PCK).
Data	Learning Forward, syllabus/template	The PD addresses school/district needs, as indicated by data.
Coherence	Desimone (2009), syllabus/template	The PD fits with what teachers already know and experience (e.g. standards).
Learning communities	Desimone (2009), Learning Forward	The PD is collaborative and involves teachers from the same context.
Active learning design	Desimone (2009), Learning Forward, syllabus/template	The PD engages teachers through active, hands-on,

		applied learning.
Duration/implementation	Desimone (2009), syllabus/ template	The PD is planned to last at least 20 hours over a semester or more, with intentional follow up.
Personalization	Syllabus/template	The PD meets different participants’ needs and provides options for them.
Technology	Syllabus/template	The PD integrates technology resources to achieve its aims.
Leadership	Learning Forward	The PD has leaders with capacity to support professional learning.
Outcomes	Learning Forward, syllabus/template	The PD uses evaluation to measure its outcomes and implementation.

The plans were coded for each feature presented on Table 1. First, I decided whether there was (a) **clear and specific** evidence of the feature, (b) **some** evidence which might lack clarity or detail, or (c) **no evidence**. The second layer of coding involved determining what type of the feature was evident. For example, for “content focused,” what was the academic content area – math, reading, etc.? Finally, I analyzed the PDP narratives qualitatively and identified quotations that either expressed common themes or particularly unique or interesting examples. My approach to analyzing the evaluation plans (i.e., the Outcomes feature) was slightly different and is explained and justified below)

Findings

This section presents categorical and qualitative data for each of the ten characteristics described above. The next section highlights implications of the findings for teacher leaders and those who develop, supervise or support them.

Table J2 Alignment with content focus feature

Evidence of feature	Number	Percentage
Clear and specific	23	65.7
Partial and/or general	1	2.9
No	9	27.3

N=33

The majority of plans focused on traditional academic content areas as shown in Figure 1, below. The remainder addressed classroom management, instructional strategies, teacher mentoring or educational technology in isolation from disciplinary or content knowledge considerations.

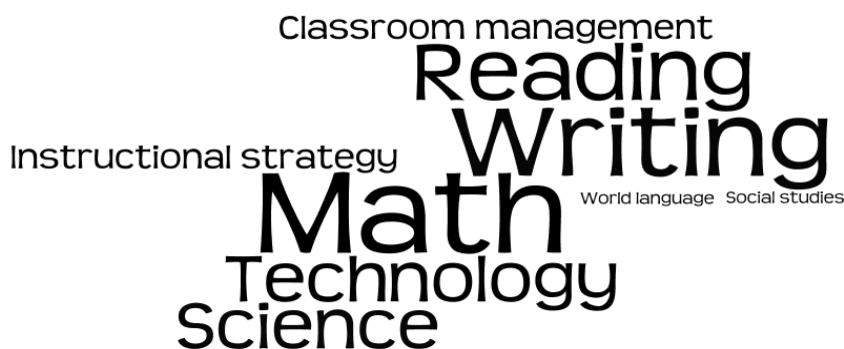


Figure J1 Content focus of PDPs

The most frequent content focus for PDPs was math, representing seven plans

(21.2%). The influence of the Common Core State Standards (CCSS) is clearly evident in these plans. Most often, PDPs focused on conceptual understanding or open-ended problem solving – skills stressed in the CCSS. Writing, also an emphasis of the CCSS, was the focus in six plans (18.2%).

Teachers described various rationales for choosing their content foci, including adjusting to new standards such as CCSS and the Next Generation Science Standards (NGSS), addressing gaps in teacher knowledge, responding to student learning needs, or implementing new programs or curricula. In general teachers were intentional about choosing their PD content although they did not often substantiate their decisions with specific data (see Table J3).

Table J3 Alignment with data feature

Evidence of feature	Number	Percentage
Clear and specific	14	42.4
Partial and/or general	8	24.2
None	11	33.3

N=33.

These ratings show relatively weak alignment on this feature, which may have to do with the format of the PDP template. In Module 1, students were asked whether and how their schools or districts use data to determine PD needs. It was intended that students would revisit and revise this section of the PD template, providing data to support their choice of a specific PD focus, but this must not have been clear. Many plans talked *in general* about how schools used data or discussed data analysis

activities that could be integrated into PD. Plans were only coded “clear and specific” if they connected the PD plans to data showing the need for that PD topic.

The plans that did feature data tended to use measures of student achievement. For example, one elementary school teacher wrote:

In thinking about discussions my team has had about district and standardized assessments such as NWEA MAP as well as formative classroom assessments and observations, there are multiple sources of evidence that reading comprehension is an area of need for professional development in my school (PDP 14)

Some plans also included teacher or administrator perception data, or teacher performance data. Several students met with colleagues while developing their PDPs to collect data first-hand.

Upon meeting with our administrators to propose our plan, they shared with us our school's biggest area of weakness. ...they have noticed trends in weakness throughout the staff in the area of “Instruction” for our DPAS II teacher observations. More specifically, the area of “3d” which...includes quality of questions, discussion techniques, and student participation (PDP 24)

Recently, teachers in my school were asked to complete the Essential 5 survey about what areas they would like additional professional development in. It was unanimous that teachers would like to have more PD around differentiated instruction. To dig a little deeper, this week I interviewed colleagues and administrators to ask what specific area [was needed]. Four out of five mentioned that they have data for each of their students and know how to analyze it, but do not know what to do with this information (PDP 55)

The feature of coherence was relatively stronger in the PDPs as a whole (Table J4).

Table J4 Alignment with coherence feature

Evidence of feature	Number	Percentage
Clear and specific	21	63.6
Partial and/or general	10	30.3
None	2	6.1

N=33.

The majority of plans situated the PD within the school, district and/or state context and gave specific evidence of consistency. For example,

As further proof of the importance of [my PDP], my principal sends a weekly bulletin to all staff members and each week it states, “Just a reminder that our school-wide expectation includes...focus on writing across content areas” (PDP 4)

Some plans also demonstrated careful thought about coherence by identifying ways in which the PD fit within prevailing expectations, and also ways in which it might conflict with or stretch those expectations. For instance, one plan, focused on writing across the curriculum, described how it reflected its school’s increasingly interdisciplinary approach but also may conflict with the district’s common assessments, in which “*comprehension skills are taught in isolation, tested and not revisited*” (PDP 14). In general, the more in-depth and complete PDPs were more likely to acknowledge the challenges associated with these ten features, rather than simply stating that they were present.

All but two plans acknowledged the importance of coherence. Ten (30.3%) made broad claims about coherence or gave examples that stretched logic. For example, one plan argued that PD on listening comprehension was consistent with a school focus on integrating technology, because technology can be used to play audio. Sometimes coherence overlapped with other features, particularly collective

participation. Some PDPs described leveraging existing collaborative structures and cultures in schools to accomplish their aims, an effective strategy.

There is already a strong collaborative connection between the different content areas within the 7th grade PLC and this strong connection is key to the success of this PDP (PDP 23)

My PDP, centered on a lesson study of a cross-grade-level lab lesson, needs a great deal of interest and energy from my peers to be successful...I anticipate quickly getting to this point given the close relationship, trust, and history the four of us have in the department (PDP 25)

All PDPs included the participation educators from the same organization.

Table J5 Alignment with collective participation feature

Evidence of feature	Number	Percentage
Clear and specific	33	100.0
Partial and/or general	0	0.0
None	0	0.0

N=33.

The specific type of teacher grouping is shown in Table J6, below.

Table J6 Types of collective participation

Types of Collective Participation	Number	Percentage
Grade-level PLC	11	33.3
Content-area PLC or academic department	9	27.3
Cross-grade or vertical groupings	6	18.2
Single grade level	6	18.2
All teachers meeting criterion (e.g., core/non-core, new, middle school, enrichment leaders)	5	15.2
Interest or readiness-based flexible groupings	5	15.2
All staff (no subgroups)	2	6.1

N=33. Many plans included more than one grouping.

The plans varied in scope, from a single grade level to a whole district (i.e., from four people to hundreds), but the latter was an exception. School (for elementary) or department (for secondary) plans were most common and plans usually grouped teachers with “like” colleagues. Many also included more than one grouping, such as grade level PLCs combined with vertical teams (e.g., K-2); interest- or readiness-based flexible groups combined with grade-level teams; or full-staff participation for some PD events followed by smaller breakout groups. Here a couple of plans describe how they plan to use collective participation and combine different collaborative groupings:

Each teacher will be on two teams, 1—grade level team and 2 – cross grade level team. The purpose of this is to provide teachers an opportunity to discuss the math and science instruction and content that is happening at their own grade level, and then talk about the instruction that is being done in each grade level [1-5] (PDP 31)

The format, which will be comprised of both data-driven discipline-specific work sessions and interdisciplinary study groups focused on the development of pedagogy to support specific skills across the curriculum, will provide opportunities for teachers to choose their areas of focus, opportunities for both interdisciplinary AND vertical learning experiences, and opportunities for teachers to build leadership experience (PDP 26)

In addition to incorporating communities of teachers, the majority of the plans featured active learning designs (Table J7, below).

Table J7 Alignment with active learning designs feature

Evidence of feature	Number	Percentage
Clear and specific	17	51.5
Partial and/or general	16	48.5
None	0	0.0

N=33.

The plans show clear attempts to engage participants in active learning and all reject “sit and get.” They differ in the amount of detail provided about those learning designs. Plans coded “clear and specific” described or explained what participants would be doing in PD and what about it would be engaging or active. Others just identify an active learning design such as a PLC or make no distinction between current practice and the PD. For example, some plans appeared to simply continue the school’s existing PLC routine. There was a fine line between leveraging existing resources (as described under “Coherence,” above) and just maintaining the status quo.

The different types of learning designs are shown in Table J8, below.

Table J8 Types of learning designs

Type of Learning Design	Number	Percentage
PLC	18	54.5
Coaching	18	54.5
Lesson study	11	33.3
Content-focused PD	9	27.3
Study group	4	12.2
Action research	3	9.1

N=33. Many plans included multiple learning designs.

Over half of the plans included PLCs or coaching. PLCs are already widespread in Delaware schools as a result of a statewide initiative under Race to the Top. Some plans described existing PLCs as well functioning while others indicated that PLCs needed to be strengthened or focused. Coaching was also a very prevalent design, with peer coaching the most frequent type. A third of the plans included lesson study, with some implementing the full process with fidelity and others adapting or

choosing parts of it. Combination approaches were common. The two examples below include thoughtful rationales for how different active learning designs fit together.

I chose PLCs because this type of PD is already established in my school by grade level and content area. There are also already teacher leaders recognized within the school and at the district level. It is necessary for PLCs to have “norms,” these have already been recognized within each PLC...Using PLCs after the initial PD session to implement and reflect on the flipped classroom will help increase its effectiveness and allow for teachers to feel successful and more comfortable (PDP 22)

[School name] already builds its schedule around grade-level content area collaborative planning but only a few groups really use the time to build and refine their instruction. We have made progress in the past by introducing elements of lesson study to the group to provide them with focus. While teachers were unable to go watch the lessons taught, they did all teach the same lessons, bring student work and – at times – recording. Layering the initial direct PD with focused procedures for this modified lesson study process for the PLC groups that will be supported by coaching oversight seems to be a nice way to bring in the new [writing] initiative while providing some correction to the [previous approaches taught through PD] (PDP 38).

As was the case with coherence, the more detailed or in-depth plans also tended to be more frank about the challenges of implementing active learning designs. These challenges include obtaining the necessary resources – especially time.

Table J9 Alignment with duration feature

Evidence of feature	Number	Percentage
Clear and specific	16	48.5
Partial and/or general	7	21.2
None	10	30.3

N=33.

The majority of plans spanned a school year; the longest scope was three years, and the shortest was a few months. As with other features, plans varied in their level of detail about how the feature of duration would be incorporated. Some plans simply stated that 20+ hours of PD would be provided but not explain when or how. Plans coded “clear and specific” identified where the PD time would come from (e.g., in-service days, designated PLC or staff meeting time, etc.). Yet, even these did not always acknowledge the trade-offs or negotiations necessary for this to happen. For instance, numerous plans intended PD to occur “during PLC time.” But what else is on the agenda during this time? Who would need to support the plan for the time to be appropriated in this way? Different levels of awareness of such issues may reflect students’ roles in their organizations.

Table J10 Alignment with follow up feature

Evidence of feature	Number	Percentage
Clear and specific	16	48.5
Partial and/or general	13	39.4
None	4	12.1

N=33.

Almost all the plans include some kind of ongoing activity indicating that participants embrace the message that “one and done” PD is insufficient. Indeed, several criticized their school’s existing PD for falling into this trap. Again, variation was mostly about the level of detail rather than the presence of the feature. Here are two examples, the first general and the second specific.

I would want to provide follow up professional development. Teachers need time to discuss what worked, did not work, and provide recommendations or strategies with their colleagues (PDP 32)

The professional development is not a one-time event. It is a series of four rounds. Each round they will gain a better understanding of the flipped classroom model and will be given multiple chances to reflect with colleagues, ask questions, and receive feedback (PDP 22).

The second design has far more structure: activities for each “round,” as well as homework assignments between the sessions (e.g., teach a flipped lesson or observe a colleague doing so) are designated. Other strategies for accountability evident in the PDPs included identifying deliverables with due dates, observing classroom instruction (by teacher leaders or administrators), or requiring teachers to bring work samples (e.g., student artifacts, assessment data) to PD sessions. These strategies were found but were not widespread. More general plans left follow-up to the teachers’ discretion, with ongoing resources or supports available to those who chose to engage.

This takes us to the feature of leadership. Who will be responsible for implementing the plan, delivering the PD and/or monitoring its translation to the classroom? This was not posed as a direct question in the PDP template, but the majority of plans provided at least some information about leadership. PDP authors who identified as holding leadership positions, either as administrators or, more often, teacher leaders, were more likely to engage this issue.

Table J11 Alignment with leadership feature

Evidence of feature	Number	Percentage
Clear and specific	18	54.5
Partial and/or general	5	15.2
None	10	30.3

N=33.

The plans demonstrate a clear preference for using “in house” PD expertise. Only five plans (15.2%) included bringing in external experts to facilitate any part of

the PD; in two cases, these individuals or groups were named, in the rest they were general ideas. More often, plans mobilized existing resources such as veteran teachers to facilitate PD. Deeper issues about capacity were only occasionally explored.

Once teachers in my school identify their area of improvement, our School Improvement Team will need to identify qualified facilitators to lead the teacher study groups (PDP 55)

Related considerations include other kinds of resources, such as time, funding, or administrative support. Plans displayed considerable variation regarding administrative involvement. Some made no mention of principals or other school leaders, while others made claims such as:

This is the backbone to having a successful learning community model in a school. The principal and/or other leadership need to be a part of the time, yet they also need to know when to step back and allow the other members to step up and take the lead (PDP 31)

I am working closely with the school's Building Leadership Team to make professional development decisions...Since the plan that I am developing targets a wide range of content, it is essential that I have leaders within the school from each content area be part of the process in order to make the professional development a meaningful experience (PDP 27. Note: written by an administrator)

We turn now to two features, differentiation and technology, which do not appear directly in Desimone (2009) or the standards but were foci of our course and are implied in research. Both features were strongly present in the PDPs.

Table J12 Alignment with differentiation feature

Evidence of feature	Number	Percentage
Clear and specific	24	72.7
Partial and/or general	9	27.3
None	0	0.0

N=33.

These PDPs clearly reject the idea that “one size fits all.” All plans acknowledged the importance of differentiation, although nine gave little or no information about how it would be accomplished. The most common differentiation strategy was to assess the needs or readiness of participants at the outset and tailor the PD to their responses. Some plans also used flexible participant groupings. Many plans distinguished roles for more experienced teachers, often using them to facilitate groups or model strategies. In general, job-embedded activities such as lesson study or coaching were seen as inherently “differentiated” because they respond to teachers’ needs and interests individually or collectively. Fewer plans provided a choice of activities; where choices exist, they tend to be somewhat bounded (e.g., teachers can choose which mini-session to attend).

Before the professional development, I would like to survey teachers for their readiness to help place them into their collaborative groups for the morning session...Teachers will also be selecting from a variety of topics for the afternoon sessions on the all-day building PD days...Lastly I will have teachers take turns leading the small groups in the follow-up after-school PD experiences (PDP 41)

Looking at the four [PBS strategies that will be a focus of the PD] I want them to pick one or two that will work best for their classroom and be a focus during the sessions. I plan to get the teachers involved by having a few lead their peers for experience. For the chill zone strategy, I have a handful of teachers that have great areas for what a [that] looks like. They can also demonstrate how it works (PDP 11)

As noted above, most PDPs were for a relatively well-defined group (Collective Participation) and demonstrated an awareness of context (Coherence). Some plans included specific information about the unique needs or demographics of their target audience.

All the PDPs infused technology to some extent and most were specific about how technology would support or enhance the goals of the PD.

Table J13 Alignment with technology feature

Evidence of feature	Number	Percentage
Clear and specific	25	75.8
Partial and/or general	8	24.2
None	0	0.0

N=33.

Nearly half of all plans (16) included Schoology, a learning management system (LMS) in use in many Delaware schools. Participants stated that Schoology was already present in their schools and teachers knew how to use it. This is also an example of coherence. Other common resources were Twitter (a social media platform) and GoogleApps (a suite of cloud-based production and collaboration tools), but there were many technologies mentioned by only a single plan.

Most often, technology was used to enable collaboration among participants. Much less frequently, collaboration extended beyond the group (e.g., to others in district or state). Other stated purposes include: (a) to promote sharing of resources (e.g., artifacts, lesson/unit plans, classroom videos, model practices); (b) deliver supplemental or differentiated content; (c) facilitate subgrouping of participants (e.g., Schoology groups); (d) provide access to non-participants (e.g., administrators); (e) to showcase work of the PD; and (f) to permit formative evaluation. Some plans proposed using Schoology posts or tweets as evidence of teacher participation, engagement, or application. Here are two different examples of clear and detailed technology infusion:

Our school will use LearnZillion, the web-based lesson planning site, during the planning stage of the lessons studies. The teachers can utilize the resources in Learn Zillion to identify their learning goals and objectives and gather ideas for planning the instructional lesson for students...By incorporating this layer of technology into our PD we are attempting to make our PD more effective by blending our online learning with our in-person learning (Dawson, 2015). LearnZillion presents an opportunity for teachers to read, watch and talk through lessons presented and they can take these ideas back into their classrooms immediately. We are also allowing teachers a chance to choose the area they want to explore in LearnZillion (PDP 31)

Schoology “subscore groups [based on SAT writing subscores] can create and share lesson plans that other teachers can use, can share data analysis and strategies for differentiating based on the data, can upload videos (either for use with students or for further teacher development) – the possibilities are endless and will ideally evolve as staff utilize the technology more and more throughout and beyond the structured professional learning event (PDP 26)

The final feature of research-based PD is a focus on outcomes and a systematic plan to measure both implementation and outcomes. (Note: this also can intersect with the feature of Data). Students were asked to design an evaluation plan for their PDP. The course textbook introduced a variety of evaluation frameworks, and almost all students incorporated at least one of these into their plans. The most common were the Concerns Based Adoption Model (26 plans, or 78.8%) and the Guskey framework (16 plans, or 48.5%) – again, the PDP template may have “led” students to these choices. Developed by researchers at SEDL, the Concerns Based Adoption Model (CBAM) is a framework for how people adopt innovations and adjust to change (SEDL, n.d.) Thomas Guskey’s evaluation model (Guskey, 2000) includes five “levels” of professional development impacts: participant reactions, participant learning, organization support and change, participants’ use of new knowledge/skills, and student learning outcomes.

To dig deeper into Outcomes, I used the Guskey framework to examine which level(s) of evaluation were evident in the PDPs and how they were measured. This is a bit more detailed than my approach to analyzing the other features (above) but still generally consistent. I am looking for evidence of the Outcomes feature, as well as insight into how participants thought about evaluating their PDPs. Figure J2 shows the percentage of PDPs evaluating each Guskey level.

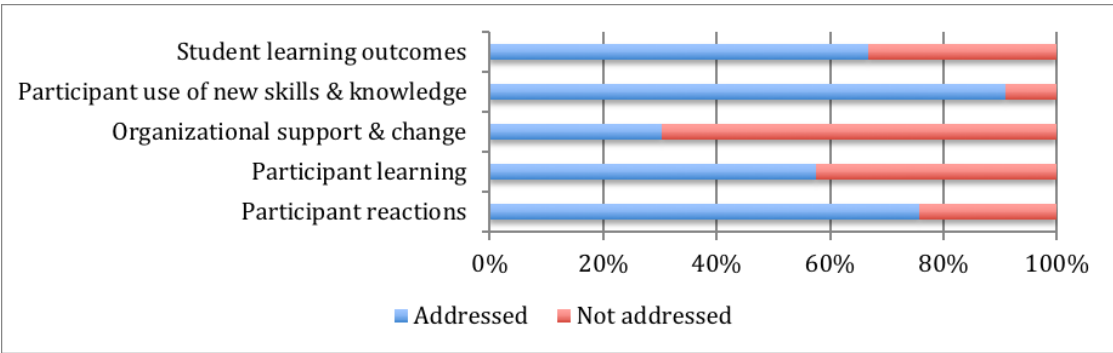


Figure J2 PDPs evaluating each Guskey level

There was considerable variation in how often the different Guskey levels were addressed in the evaluation plans. All but three PDPs (90.9%) planned to assess participants’ *use* of new skills or knowledge (i.e., teachers’ instruction) but only 19 (57.6%) directly measured participants’ learning. Notably, only ten (30.3%) included measures of organizational support or change.

Students proposed a range of methods to gather data about the implementation and/or outcomes of their PD. Table J14 cross-references these methods with the level of evaluation at which each would be used.

Table J14 Evaluation methods by Guskey level evaluated

	Level 1: Participants' reactions	Level 2: Participants' learning	Level 3: Organizational support/change	Level 4: Use of new knowledge/ skills	Level 5: Student learning outcomes
Survey/exit ticket	23	10	2	10	2
Interview/check in	9	6	-	-	1
Observation/ Walkthrough	-	1	2	20	-
Artifact/ Document review	-	-	2	5	2
Assessment data	-	1	-	-	16
States level will be measured, not how	-	4	4	-	1
Other	-	-	-	5 (Twitter, LMS) 1 (mandated)	1 (climate data)
Not addressed	8	14	23	3	11

N=33. Many plans included more than one data collection method for each level evaluated.

Most plans incorporated feedback from participants, especially through surveys or questionnaires to gauge PD satisfaction, learning, and application. Interviews were also fairly common; these often were framed as informal “check ins” rather than structured processes. Beyond teacher self-reports, most plans also included observations or walkthroughs to examine application of PD to the classroom. A few plans also used other methods such as reviews of documents (e.g., PLC minutes) or technology.

PDPs described a range of evaluation purposes. Most often, they planned to use findings to improve the current PD or to plan or differentiate future PD. PDPs often framed evaluation as a way for participants, usually teachers, to express their views and their needs so that the PD could better meet them. A few PDPs demonstrated a broader perspective, such as suggesting that evaluation processes or results could inform stakeholders, build “buy in” and momentum for change, or create

institutional memory. Only a few PDPs focused on summative uses of evaluation (e.g., determining whether the PD was successful, reporting for accountability, etc.).

In general, plans did not address issues of evaluation capacity, such as who would conduct evaluation activities or whether those individuals would have the time or technical skills to do so. None of the PDPs proposed involving external evaluators. Connections between evaluation questions, data collection activities, and uses of findings were often not explicit. These PDP analyses align with our sense as instructors that students could build only a basic understanding of PD evaluation in a week, especially the last week of the semester.

Implications

In this artifact, I have reviewed 33 PDPs using a structured framework built around characteristics of effective PD. Since EDUC 774 was designed to reflect research about PD, it is perhaps not surprising to find evidence of most of the features, in most of the plans. Variation lies more in the level of depth and detail with which the feature is discussed, rather than its mere presence. The format of the PDP template also strongly influenced these findings. Almost all students chose to use this optional template, which “chunked” the PDP into specific sections, aligned with weekly topics in the course which were in turn aligned with major areas of PD research (e.g., content focus, evaluation). Two of the features in my framework did not explicitly appear as sections in the PDP template: data and leadership. These were the two features most likely to be omitted from the plans. We may want to pose more direct questions about data and leadership in future templates. We must also acknowledge the limitations of these data, since the assignment shaped the way students presented their ideas about professional development.

Generally, my analysis suggests an orientation towards engaging, ongoing, and relevant professional development. Relevance was generally achieved by rooting the PD as “close to home” as possible, such as within the same content area or grade level. Often, PDPs included classroom-embedded activities such as coaching. There was some interest in making connections across grade levels within a single school (i.e., vertical articulation) but connections across schools or districts, or collaborations within a broader professional community were much less common. The PDPs generally attempted to address teachers’ specific and differentiated needs.

These plans demonstrate a strong interest in leveraging teacher expertise from within the school. This is consistent with the emphasis of EDUC 774 and indeed the Teacher Leadership program as a whole. Many PDPs called upon teachers to model practices, facilitate sessions, coach their peers, or oversee the entire PD effort. These strategies provide leadership opportunities for teachers and contribute to the relevance and credibility of the PD. At the same time, this raises questions about capacity and resources. Does the existing staff have the knowledge and skill to provide the PD? What about the *time* to develop, deliver, and evaluate that PD? Some plans addressed capacity questions but many did not. Relatedly, they varied in the level of detail and apparent understanding of what it takes to fully implement a PD initiative. In many cases, the plans seemed aspirational and/or ambitious; they would require substantial problem solving to be implemented in real schools. Through my analysis, I realized that our PDP template, and indeed the course, did not explicitly address the Learning Forward standard of Resources. Highlighting this standard in the future may help us push students to confront these questions and develop more feasible plans. We may

also want to help students see the value of external perspectives (e.g., from outside presenters, facilitators, or evaluators), especially to stimulate change.

Taken as a whole, the PDPs are somewhat optimistic about teachers' willingness and ability to adopt new practices. Following the Guskey framework (Guskey, 2000), the plans tend to assume that if the PD is well designed and engaging, and participants are satisfied and learn new skills (Levels 1 and 2) that they will then use these new skills (Level 4) and students will benefit (Level 5). The plans' lack of attention to Level 3 is notable in this context. Few seriously grappled with the organizational factors that might either facilitate or impede PD implementation. For instance, although many PDPs stated that a long duration (20+ hours) was important, few specifically considered how that time could be negotiated. Many PDPs included observations or walkthroughs to check whether PD was being translated into classroom practice, but few explored the broader conditions, such as accountability or trust, that might encourage this to happen. Students seemed to have varied grasps of the complexity of instructional or organizational change, probably reflecting different roles or levels of experience.

Some held school or district leadership positions that provided them with broader perspectives, but many were more restricted to their own classroom experience. These students were themselves motivated enough to join a master's program in teacher leadership, and skilled enough to be accepted. Their PDPs sometimes seemed to assume that the audience of fellow teachers would be equally engaged. It is not clear that many of these PDPs would move a teacher who did not independently see the need to change.

Finally, these plans suggest that Delaware public schools have some PD strengths in place. While most students acknowledged the need to improve PD in their schools, only a few proposed a very large or “second order” change and/or strongly critiqued the status quo. Indeed, many PDPs talked about incorporating existing resources or structures – ranging from Schoology to new curricula, from PLCs to teacher leadership roles – into their plans for the future. Coherence was a strength of these PDPs. It appears that many teachers felt their school or district contexts had foundations upon which they could build more effective PD.

References

- Desimone, L. (2009). Improving impact studies of teachers’ professional development: Towards better conceptualizations and measures. *Educational Researcher*, 38 (3), 181 – 199.
- Guskey, T. R. (2000). *Evaluating professional development*. Newbury Park, CA: Corwin Press.
- Learning Forward (2011). Standards for professional development. Retrieved April 21, 2014 from <http://learningforward.org/standards>
- SEDL (n.d.). The Concerns Based Adoption Model. Retrieved February 9, 2017 from <http://www.sedl.org/cbam/>

Appendix K

ARTIFACT K: EXPLORATION OF EVALUATION USE WITH FORMER CLIENTS

Introduction and Method

The purpose of this artifact is to interrogate one of the central assumptions of my ELP: that research and evaluation can improve professional learning opportunities. Universities have several levers they can use to effect changes in K-12 education. One of these is to conduct evaluations of professional learning programs, which then can be used to create stronger, more research-aligned designs, implementations, or leadership. Taking advantage of my prior experiences, I sought to learn more about whether, and in what ways, this occurred for four professional learning programs I evaluated between 2012 - 15 as a staff member at the Delaware Education Research & Development Center (DERDC) at the University of Delaware (UD).

Within this broad overall purpose, this inquiry has several other goals. I wanted to build awareness of what professional learning practitioners value in and want from program evaluations. Better understanding their perspective and needs could help me improve as an evaluator. I could also distill insights from this inquiry to inform colleagues within the university or even beyond. I wanted to try a new approach to obtaining feedback from clients. Finally, as I continue to hone my professional goals, I hoped that this process would teach me more about program evaluation and professional learning and how they intersect. I therefore sought to answer the following seven questions:

1. What kinds of programs were evaluated, and how were the evaluations designed?

2. How did clients understand the purpose of the evaluation? What did they expect?
3. In what ways were the evaluations used?
4. What factors contributed to evaluation use, or detracted from it?
5. What did clients most value/appreciate about the evaluation and my work with it?
6. What suggestions for improvement did they offer?
7. What are the implications of these findings for my own understanding of program evaluation and its link to improving professional learning?

Before reviewing my methodology and answering these questions, I provide a brief overview of the literature on evaluation use.

The basics of evaluation use

Evaluation is a broad field. Alkin (2004) developed a metaphor for understanding its development, a “family tree” with three main branches, each representing a priority: use, methods, and valuing. I align myself with the “use” branch. I came to evaluation from teaching and non-profit leadership and I see the goal of my efforts as improving educational programs. For improvement to occur, use must occur – but what does that mean, and how does it happen? There is a large literature on evaluation use, and I acknowledge I have only dipped into it. In this section I will examine three main types of evaluation use: instrumental, conceptual, and process.³

Instrumental and conceptual both refer to the use of evaluation findings, but the former are more concrete and specific applications. There are many forms of

³ A fourth category in the literature is political use, also sometimes called persuasive or symbolic (Weiss, 1998). This involves using an evaluation’s findings, or its mere existence to justify previously held positions. Since this is not a type of use to which I aspire, I did not focus my data collection on it.

instrumental use; one way to categorize them is provided by the evaluation use sub-scale from the Evaluation Capacity Assessment Instrument (ECAI) (Taylor-Ritzler, Suarez-Balcazar, Garcia-Iriarte, Henry & Balcazar, 2013). Uses include:

1. To report to a funder.
2. To improve services or programs.
3. To get additional funding.
4. To design ongoing monitoring processes.
5. To assess implementation of a program.
6. To assess quality of a program.
7. To improve outreach.
8. To make informed decisions.
9. To train staff.
10. To develop best practices.
11. To eliminate unneeded services or programs.

Each of these uses implies action and/or decision-making. By contrast, conceptual use is broader and involves changes in thinking. As Carol Weiss (1998) writes, “even if [program staff] are blocked from applying the findings to decisions at the time the study is reported, the findings can change their understanding of what the program is and does. They gain new ideas and insights” (p. 24). For example, evaluation may help staff recognize their program’s strengths and weaknesses. Finally, process use refers to what stakeholders learn by being involved in the evaluation rather than by its results or findings. Patton (2008) defines these as: “individual changes in thinking, attitudes, and behavior, and program or organizational changes in procedures

and culture” (p. 155). For example, staff may learn how to set more specific goals for their program.

Evaluation use is highly contextual. Early research sought to correlate evaluation features to instances of use, but now it is generally accepted that use “pathways” are more complex than discrete variables (Johnson et al, 2009). Still, research provides some guidance about the circumstances in which evaluations are most likely to be used. Participation of stakeholders in the evaluation promotes use, as does timely and clear communication. As Weiss states: “The best way that we know to date of encouraging use of evaluation is through involving potential users in defining the study and helping to interpret results, and through reporting results to them regularly while the study is in progress” (p. 30). Similarly a 2009 literature review of evaluation use “point[ed] to the importance of stakeholder involvement in facilitating evaluation use and suggest[ed] that engagement, interaction, and communication between evaluation clients and evaluators is critical to the meaningful use of evaluations” (Johnson et al., 2009, p. 389). Communication that takes place over time is especially powerful (Huberman, 1989). Evaluations are more likely to be used if they are seen as credible and relevant. Finally, Patton (2008) elaborated a concept he called “the personal factor” – people or groups within an organization who care about the evaluation and are its primary users.

The personal factor represents the leadership, interest, enthusiasm, determination, commitment, assertiveness, and caring of specific, individual people. These are people who actively seek information to learn, make judgments, get better at what they do, and reduce decision uncertainties (p. 66).

These research findings are reflected in professional standards as well. The Joint Committee on Standards for Educational Evaluation (2011) established eight

“utility” standards, which are “intended to increase the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs.

U1 Evaluator Credibility Evaluations should be conducted by qualified people who establish and maintain credibility in the evaluation context.

U2 Attention to Stakeholders Evaluations should devote attention to the full range of individuals and groups invested in the program and affected by its evaluation.

U3 Negotiated Purposes Evaluation purposes should be identified and continually negotiated based on the needs of stakeholders.

U4 Explicit Values Evaluations should clarify and specify the individual and cultural values underpinning purposes, processes, and judgments.

U5 Relevant Information Evaluation information should serve the identified and emergent needs of stakeholders.

U6 Meaningful Processes and Products Evaluations should construct activities, descriptions, and judgments in ways that encourage participants to rediscover, reinterpret, or revise their understandings and behaviors.

U7 Timely and Appropriate Communicating and Reporting Evaluations should attend to the continuing information needs of their multiple audiences.

U8 Concern for Consequences and Influence Evaluations should promote responsible and adaptive use while guarding against unintended negative consequences and misuse.

Instrument

These standards have been incorporated into instruments and processes to design useful evaluations or assess the utility of an evaluation. I drew on some of these to develop my interview protocol for former clients. For instance, Stufflebeam (1999) developed a checklist for meta-evaluation, i.e., reviewing evaluations to see to what

extent they demonstrate the professional standards. I also examined Patton's (1997; 2004) much more robust model of utilization-focused evaluation (UFE) and associated tools, such as a checklist for planning a UFE (2013). Finally, I consulted the literature for tools for obtaining feedback from evaluation clients. Although I did not find any interview protocols, I located one client feedback form developed by a Topical Interest Group of the American Evaluation Association (Doyno-Ingersoll, Haley, Dowell & Chambliss, 2005). I used these tools along with the ECAI (Taylor-Ritzler, Suarez-Balcazar, Garcia-Iriarte, Henry & Balcazar, 2013) to help me develop interview questions.

It is important to acknowledge that I did not apply these tools as intended but rather used them more broadly to inform my thinking. My rationale for using interviews rather than surveys was to obtain in-depth insights and be able to ask follow-up questions. Since I was no longer in any official role with the programs, I anticipated high levels of candor. This might have been more challenging had this interview occurred *during* the evaluation. At the same time, since time had passed I recognized that clients might need clarification or refreshers about the evaluation. Interviews provided the best opportunities for such dialogue.

My interview protocol is in Appendix 1. The instrument crosswalks interview questions with dimensions of evaluation use derived from the literature and standards discussed above. These included dimensions related to:

- Context (e.g., project background, political climate, “personal factor”)
- Stakeholder involvement
- Evaluation processes (e.g., questions asked/answered, methods)
- Evaluation products (e.g., relevance, clarity and timeliness of reporting)

- Use of findings (conceptual, instrumental, process)
- Evaluator characteristics (e.g., competence, credibility, interpersonal)

Sample

I developed my sample based on five criteria, approaching clients who:

1. Managed a project or a substantial piece of a project;
2. Focused on professional learning for educators (K-12, higher education, or informal);
3. For which DERDC was the external evaluator; and
4. I served as the Principal Investigator (PI) or key/sole contact for the evaluation; and
5. With whom I had ongoing communication since 2015.

The first four criteria were straightforward. I added the fifth assuming that I would have a better chance of securing participation from people with whom I had some relationship, though this may also influence my data. Five individuals met these criteria. All initially agreed to participate but I was not able to schedule an interview with one, making my final response rate 80%. This is a small sample with inherent limitations, but it reflects a cross-section of the types of programs I evaluated (see question 1, below). Of my participants, two were University of Delaware (UD) professional staff members and two were UD faculty. Just one was the overall PI for the project I evaluated; the others managed specific work groups or coordinated the project.

Method

This project was submitted to UD's Institutional Review Board (IRB) and approved. I contacted potential participants over email. Once they agreed to participate

and an interview was scheduled, I sent them preparation materials including: (1) a short (2-3 page) summary of the evaluation, including evaluation questions, data collection methods, communication between program staff and evaluator, etc. and (2) one or more evaluation reports from 2014 or 2015. I requested they review these materials to refresh their memories before our conversation. Interviews took place in April 2017 either in person or over Skype averaging approximately 40 minutes. Interviews were digitally recorded, transcribed, and analyzed qualitatively for themes using Dedoose software (SocioCultural Research Consultants, 2016). The coding framework included 16 codes and followed the protocol dimensions listed above.

Findings

What kinds of programs were evaluated, and how were the evaluations designed?

The four participants represented three funded grants. Although two participants were technically within the same grant, they were responsible for separate activities and did not collaborate, so I consider them separate “programs.” All were professional learning programs for educators with varied content foci including science, technology, engineering and/or math (STEM) (3), English Language Arts (ELA) (1), and social-emotional learning (SEL) (1). Some programs covered more than one content area. Program designs and target audiences also varied. The most common activity was workshops (half-day, full-day or multi-day) (3) followed by one-on-one coaching (2). One program also featured curriculum (re)design activities and one included university coursework. Two programs targeted in-service teachers in Delaware K-12 schools. One was an alternative route to certification, reaching individuals from the summer before they began teaching through their first few years

in the profession. Finally, one program provided activities both for higher education faculty and for pre-service teachers. In addition to designing and leading activities, in two cases program staff members also conducted research on professional learning.

The programs also varied in scope and structure. Two were funded by five-year federal grants, and I acted as an evaluator for both from very close to the start of year 1. One of these had a small budget (approximately \$300K), while the other was much larger (\$5M) with multiple working groups. For the latter, I worked under the direction of a more senior evaluator. The third grant was a one-year, state-funded competition that supported multiple concurrent professional learning activities (of which two are represented in this study). UD received this grant for three consecutive years, but activities varied. I was the PI for this project and supervised several graduate research students. Table K1 lists participants and the projects with which they were involved.

Table K1 Participants and projects in the study

Participant	Project	Content Focus
A	Five-year, \$300 K project funded by the US Department of Education	Pre-service teacher education in ELA, Math, Science, Technology
B	Five-year, \$5M project funded by the National Science Foundation	Science, Engineering
C	One-year, \$200 K project funded by the Delaware Department of Education	Social/Emotional Learning
D	One-year, \$200 K project funded by the Delaware Department of Education (same project as Participant C, but different set of activities)	Engineering, Computer Science

The scope and length of the projects influenced the evaluation design. All were based on project goals and objectives, organized around a series of questions, and more or less explicitly aligned with the Guskey (2000) framework for evaluating professional learning. In this model, professional learning has five “levels” of impact: participants’ response, participants’ learning, organizational support and change, participants’ use of new knowledge and skills, and student learning outcomes. An evaluation should measure or inquire into as many of those levels as feasible; higher levels are more meaningful but take more time and resources to evaluate.

All four evaluations studied in this artifact measured participant responses to and learning in the activity, as well as whether it reached the target audience (Levels 1 and 2). We also looked at measures of organizational and instructional change (Levels 3 and 4) though this was very limited in one-year projects. Measures of teacher or student outcomes (Level 5) such as DPAS II data or school climate data were included in two evaluations. All the evaluations were mixed methods and included surveys of participants, review of project records, observations of program activities, and review of participant artifacts. All but one included follow-up interviews with participants and others (e.g., supervisors, coaches) but the depth and range of the data collection depended on the project. There were also some evaluation activities that were unique to one project, such as pre/post content testing.

Weiss (1998) argues that evaluation approaches form a spectrum, with traditional “evaluator in charge” at one end and empowerment evaluation, in which stakeholders conduct their own evaluation, with minimal guidance, at the other. All four evaluations would fall between these extremes, with a longer timeframe allowing for more participation. In all cases, clients had the chance to provide feedback on

instruments and procedures. In the longer projects and/or where project staff was conducting their own research project, we had more extensive conversations about evaluation questions and data collection methods. None of the evaluations featured deeply participatory strategies (e.g., co-interpreting data).

How did clients understand the purpose of the evaluation? What did they expect?

The answer to this question depended on clients' prior experiences with evaluation and developed over time. Only one was directly involved in the selection of DERDC as an external evaluator. S/he had worked with us in the past and saw the choice as straightforward. All others were essentially informed by their PIs that there would be an evaluation, and that DERDC would conduct it. One initial purpose for planning an evaluation was to get funded:

It was a requirement, and actually a big portion of the proposal. I think it's something like 20 percent of the score relies on your evaluation.
(Participant D)

Thus, beginning with the proposal phase, the program and its evaluation – and the quality and even feasibility of both – were intertwined. Once funding was obtained and the program was launched, clients expected the evaluation to satisfy funder requirements. All four described this accountability purpose, a form of instrumental use as indicated in the ECAI:

We incorporated [evaluation reports] into the annual performance reports for Washington. (Participant A)

In most cases, participants' original expectations for the evaluation did not go far beyond this instrumental use. Two had never worked with external evaluators and felt some trepidation at the prospect.

Well, I was very suspicious at first. I had never worked with evaluators and I did not really know what the role was. (Participant B)

Reflecting back, both described how far their knowledge about evaluation had developed. As time progressed, they became more engaged and invested. A third participant had very limited prior experience and s/he too described deepening appreciation for what evaluation could do.

You help everybody to understand why there was value in you having a role in that work. So first you said to me, “It is a requirement, we have to do this,” but then as you got into it and you sought to understand what we were doing— and didn’t just take this on as “This is my job” — you really listened and said, “So what are the questions? What are the key things that you would like to know about?” And you framed those into what you did, and you took the time to go out and observe the sessions and really get much more familiar with it. You really became part of the team, right? (Participant C)

The interchange recounted here is an example of standard U3 (“Negotiated Purposes”) in action. The fourth participant had deeper understanding about evaluation and the contributions it could make to project management and research. S/he had previously worked with DERDC as well as other evaluators. Client attitudes and experiences are discussed further under question 4, below.

In what ways were the evaluations used?

All four evaluations were used, but in different ways and to different extents. I coded my data looking for various kinds of instrumental, conceptual and process uses and found mostly instrumental examples. As described above, the fundamental instrumental use of evaluation to satisfy a funder and comply with accountability expectations came up in all four interviews.

Beyond this purpose, one evaluation appears to have been used in only one additional way: to supervise staff members, an instrumental use (Taylor-Ritzler et al.,

2013). In this case, the evaluation confirmed staff performance issues, which program managers could then address.

I think you and I had some off-line conversations about those things that were happening. We were able to kind of tee that up so we could help [staff member]....

I'm just pretty open and frank and I would just say, "[Name], I heard this and I just need you to know that that's a problem and we've got to figure it out. What do you need me to do to help you figure it out?"
(Participant C)

This client praised my “*grace and humility*” in handling a potentially awkward situation and recognizing its impact (Standard U8, “Concern for consequences and influence”). This was the only additional use of this evaluation and conversely the only detailed example of this type of use. Barriers to evaluation use in this situation, and in general, are discussed below.

The three other evaluations were applied in many more – and more varied – ways. Most often the examples wove together two instrumental purposes from the ECAI: “to improve services or programs” and “to make informed decisions.” Clients described using evaluation findings for both concrete long-term/annual improvements and shorter-term adjustments. One program restructured its admissions process and training timeline based on Year 1 experiences (included but not limited to the evaluation). Another client described how a prior evaluation report helped the team better understand participant needs and re-allocate time and resources to meet those needs.

The evaluation report from the year before really informed how we designed this particular summer PD. So we've added those three sessions for skill base because we thought it was too much – one of the things I think that came out of the previous work was always like they needed more time...So we tried to space things out in terms of skills,

and also they really, really had a hard time making the connection to their classrooms. (Participant D)

By putting more emphasis on classroom connections in Year 2, this program increased its coherence – a research-supported feature of effective professional learning. My colleagues and I tried to be explicit about these features, for instance by referencing the Standards for Professional Learning (Learning Forward, 2011) formally in data collection instruments and informally in client conversations. I see this as an example of Standard U4 (“Explicit Values”). There were other examples where evaluation efforts helped programs better align with research and best practices. One client described re-designing a workshop to include more hands-on activities, scaffolding, and relevant projects, and less content. In this case the stimulus for change was both an evaluation report and ongoing conversations between the evaluation and project teams about professional learning. These conversations helped the project staff, new to designing workshops for adults, develop their understanding of “what works” and why. In this way, the evaluation had both instrumental and conceptual uses. It led to concrete change and also helped clients view their program through a new lens.

More generally, examples of decisions based upon evaluation findings include: increasing budget/time resources for certain parts of the program (2 examples), revising schedules (2), curricula (2) or admissions processes (1), changing service areas (1) and determining whether or not to continue an activity (2). These are all examples of instrumental use.

Since I observed program activities, another use of the evaluation was highly formative. Three clients spoke about how they used my informal feedback to make quick adjustments between days or sessions – or even within them.

We are asking for feedback on the fly – taking the temperature of the people who are participating, and thinking about how things have gone

so far, and what is the best way to use the time we have left to reach our outcomes?

Interviewer: Is that feedback useful to you guys?

Super useful, really really helpful and I think kept us from boring [participants] to tears. (Participant B)

In such situations, an evaluator functioned as embedded eyes and ears. Some clients implied that this feedback was equally or more valuable than my formal communications; I had not realized how much they valued it. One client also described ongoing communication with me as a form of monitoring. The ECAI describes a potential instrumental use of evaluation “*to design ongoing monitoring processes.*” My example seems rather to suggest that evaluation was the process:

The purpose of the evaluation was to make sure one – that we were fulfilling all the parts of the grant; two – to keep us on track; and three – to provide us with a mirror so we could assess our performance and improve where needed.

Interviewer: Okay, what does it mean – the second thing – to keep on track?

Fulfill all of the main objectives. Year 2 you should be doing this. And we had those conversations – “we’re not doing this.” “Well, why aren’t we doing it? Is that valid?” And there was a lot of back and forth with that. (Participant A)

Two projects had research components; as we were conducting the evaluation, other staff members were conducting parallel research about professional learning. I contributed data and analysis to their efforts and was listed as a co-author on several presentations and manuscripts. This somewhat aligns with the ECAI’s category of evaluation used “to improve outreach” -- or in this case scholarly dissemination.

That was a pretty important component and a critical one because basically all of our findings came from the [evaluation] report except from the pre and post open-ended questions. (Participant D)

The final item on the ECAI sub-scale is “to eliminate unneeded programs or services.” I found no examples of full elimination, but two evaluations were used to bound program activities. In one case this entailed limiting the number of partners and activities written into a proposal for future funding; in another, it meant narrowing the target audience and the program goals.

Especially in the second situation, the narrowing seemed to reflect both instrumental and conceptual uses (i.e., actions the program staff took, as well as changes in their understanding of what the program was about). This client described how the evaluation helped him/her to develop more concrete goals – “*to take them away from pie in the sky and make them action items*” – and avoid mission drift – “*to become attuned to focus back in. Five years is a long time.*” In this same project, evaluation highlighted the breadth of the work by identifying how differently participants were approaching a scientific topic. Our findings helped the project staff acknowledge:

[It is] very large and broad topic, which is great because there are lots of opportunities for a point of entry. However, that means that people are going to take different pieces of a very large pie. So it makes it very complicated. (Participant B)

Although two clients described learning about evaluation and gaining appreciation for its value, only one gave solid examples of process use. S/he felt the team “*learned a lot how to collect data and conduct research,*” specifically how to design surveys and ask for feedback in an open-ended, non-leading way. Ironically this individual also described how the evaluation also helped him/her realize “*the data isn’t everything*” and it was also important to “*trust your judgment.*” This client described him/herself as a very quantitative person who came to appreciate more qualitative considerations and exercise professional discretion.

What factors contributed to evaluation use, or detracted from it?

Facilitators

As suggested in the literature and utility standards, timely and meaningful reporting promoted use (Standard U7). Participants generally felt they had the information they needed, when they needed it. This required prioritizing based on intended use (Standard U5, Relevant Information). For instance, in one project we agreed to turn around workshop reports very quickly, so changes could be made month-to-month, but to move slower with annual reports whose main purpose was to submit to the funder. Participants stated that reports were generally effective (Standard U6, Meaningful processes and products):

I found that they were written in a language I could understand even though it's not my area of expertise. (Participant B)

Most reports ended with recommendations for the future, which participants said they found useful; one said s/he usually flipped straight to that section. In general, recommendations did not contain surprises or major *ahas*. Rather, they confirmed and crystallized ideas that were already under discussion.

Indeed, I learned that reports were necessary but insufficient for facilitating use. Ongoing discussion in the context of a trusting relationship between project staff and the evaluator(s) seemed more important. When participants described what they used (see previous section), more often it was an insight from a conversation with me rather than a report. Client trust develops over time and was strongest in multi-year projects and/or where there were existing relationships. Without prompting, two clients contrasted where our collaboration began and where it ended. Both were new to managing projects and working with evaluators; they saw reciprocity with my learning curve as an evaluator:

You were working on your style, and I was working on mine....we learned together. (Participant A)

In this case, my newness to the profession did not seem to compromise my credibility as an evaluator (Standard U1). Interpersonal skills are discussed further in the next question.

Regular communication facilitated evaluation use. Of course, this required resources of time and therefore funding, for both the program and its evaluation. There was something of a “virtuous cycle” at play: the more we communicated, the more clients used the evaluation and saw its value, and thus renewed their engagement. Again, this occurred in the multi-year project much more than the one-year projects.

Finally, some participants experienced encouragement from others to use the evaluation. This was strongest in a multi-year project whose federal funder provided positive feedback about the evaluation and the way project staff made adjustments based on data. The political context here was very conducive to evaluation use:

After Year 1 we called Washington to say we want to make some changes, and there was the realization of Washington saying, we expect you to make changes based on data. If you’ve got data ... do it.

The stuff that we couldn’t get to we were able to justify with Washington why we couldn’t do it because of the back and forth. And Washington...they told [the PI] this is one of their success stories because of the relationships and the stuff that’s happened. And a lot it’s because of the back and forth we had [with the evaluation team]. (Participant A)

Buoyed by this support, this client developed a strong investment in the evaluation and s/he helped others buy in as well. S/he took time to introduce me to participants and explain the purpose of the evaluation, which helped with response rates:

In the beginning [I would] say, speak your mind; it's why [the evaluators] are here. This is what we do. And also the big thing was, too, in August to come in and say, based on your feedback from July here's what we're doing differently. And I think they took it seriously because of that. (Participant A)

Another example of Patton's "personal factor" was the client who acted as a liaison between the evaluation team and her project team. As she grew increasingly comfortable with the evaluation, s/he attempted to persuade colleagues of its value though s/he continued to meet some resistance. For instance, s/he would inform others on the team about upcoming evaluation activities and solicit their input. To address colleague concerns (for instance, about the evaluation scope) s/he tried to communicate what was involved with collecting and analyzing qualitative data. However, s/he stated, *"I think my crew remain unaware of how much work it is"* (Participant B). This observation leads to the next section, about detractors to evaluation use.

Detractors

This section focuses on two barriers to evaluation use: time and lack of understanding or communication related to the evaluation. My data demonstrate that there is less use in a one-year project. A short grant period makes it difficult to build the relationships that facilitate trust and use, as described above. Moreover, both project staff and evaluators are likely to have less time allocated to such projects. This means there are fewer opportunities to discuss, reflect on, and make decisions based on evaluation findings or processes. Here a client reflects on "what could have been" in a longer or larger project:

It really didn't fit totally with my job, right? And so, I was doing many things at one time and making that— If I could've just done this all the time, it would've been great. ...So, then your use in it, your involvement

in it would have been a completely different thing, and I would have called on you even more indifferently had I just been able to devote more time and just had your aid along the way. (Participant C)

Relatedly, short projects limit the instructional changes that can be reasonably expected and thus the depth of the evaluation questions that can be pursued. Clients who were well versed in professional learning recognized the limits of their projects and, correspondingly, the evaluations.

These are the important kinds of [evaluation] questions. I mean, there's nothing that's not important about these; just—can you answer them in a year? A month? A grant that's so time constrained and trying to get stuff done. (Participant C)

They saw the evaluations as necessarily somewhat superficial. They understood why this was the case, but they would have found more value, and potentially more use, in data collected over a longer timeframe that examined changes to teaching and learning in more depth. That would have generated more relevant and meaningful information (Standards U5 and 6). This idea is further explored in questions 5 and 6, below.

Other detractors were organizational and cultural. If encouragement from supervisors or funders facilitated evaluation use, the reverse was also true. When there was a lack of investment from PIs and funders, or when there were transitions or administrative changes, utility suffered. Also, project changes could come down “from above.” In the situation below, an activity was organizationally re-located from one center to another, unrelated to the evaluation:

That's the loop we didn't really close on this one, so you wrote your report— the world went on, [staff member] went on— I went on and [administrator] came over and said, “No more.” (Participant C)

In two projects, there was also resistance to the evaluation. As mentioned above, two clients, both new to evaluation, initially did not understand the role of an

external evaluator, how we could add value or what engagement between project and evaluation staff would look like.

Both described how their own comfort and buy-in developed.

It was an evolution because in the beginning I didn't know what to do. And I kept you at arm's length. And then it evolved over time where we – I can only speak for me. I started trusting you. (Participant A)

However, their colleagues did not necessarily build the same trust. Project structure became a barrier to evaluation use. In large project teams/working groups, it seemed more efficient for just the leader to interact with the evaluation team and then bring back information and insights to the rest of the team. However, this could also lead to misunderstandings or missed opportunities to build relationships. My attendance at group meetings seemed to facilitate use but was not always possible. Coordination and communication issues may be inherent in large projects, but the added discomfort related to evaluation can magnify them. Strategies for how I could have handled this situation are discussed in question 6, below.

Finally, dissemination of information was more challenging in larger, less coordinated projects. In interviews, I learned that one client had never seen one of the reports from his/her project, and another could not remember if s/he had. Neither one saw this as a problem; they were familiar with the information in the report (i.e., from conversations with me and/or earlier reporting). Sometimes the question of to whom to send a final report is not straightforward (i.e., do you send it just to the PIs, and let them decide whether/how to share?). Still, I see this as a missed opportunity for communication and possibly for use (Standard U6, Meaningful products).

What did clients most value/appreciate about the evaluation and my work with it?

My interview protocol focused on the “evaluation” and I tried to frame it broadly to include the evaluation plan/questions, data collection methods, analysis, and reporting as well as ongoing and informal communications. This was to emphasize that we were talking about much more than a report. Because of this breadth, at times the “evaluation” (what we were doing) and the “evaluator” (who was doing it) became conflated in the data. Also, these data are situated. They suggest what clients preferred for a specific program context, given an assigned evaluator. It may be that under different circumstances (e.g., a larger number of participants) their preferences would be different.

All that said, clients seemed most to value data that they could not have collected themselves, whether because of the time, skills or stance required. For instance, one client commented:

The [evaluation questions] that are important to me are all the ones where we're getting feedback from either the [teachers] or the coaches or the school. (Participant A)

S/he also felt that as an external evaluator I could potentially obtain more candid, credible feedback. A second client described the evaluation methods s/he found most meaningful:

I think it's kind of the one-on-one where you can read a person and get a sense of what they're saying, what they're thinking – tells you a lot more. Or a focus group kind of thing. Told you a whole lot more than it does when you just get a bunch of papers. (Participant C)

This comment suggests a preference for interviews or focus groups rather than surveys (“a bunch of papers”). Since qualitative data collection and analysis require resources (time, training), clients saw this as a contribution the evaluation team could potentially make. They expressed other “qualitative” interests, such as enjoying

reading quotations from participants or rich descriptions of program context. Again, this finding may reflect the nature or size of the programs and/or my own orientation and skill set. In terms of evaluation reporting, clients spoke with more enthusiasm about ongoing communication rather than formal reports. As described above, most examples of use seemed to be tied to conversations with the evaluation team rather than to discoveries from a report.

I also sought direct feedback on my skills and my role within the program. Despite my lack of experience in program evaluation, I apparently had sufficient credibility. Some of this I gained second-hand, through my center and my supervisor:

I had confidence. [Supervisor] had said you were the right person and so I trusted her.(Participant B)

Some clients knew my professional background, which helped. They also saw me establishing credibility and rapport with project staff and participants, especially K-12 teachers. Two clients used similar phrasing to discuss how I positioned myself interpersonally:

You struck a very nice balance between engaging people and staying separated from them.... So, you could engage with people, yet be credible and professional. You didn't go too far. Hanging out with them would have compromised the data – I really felt like ... people were comfortable talking to you, which was really, really nice. And I have a feeling they were comfortable speaking their mind. (Participant A)

I think you've done a great job establishing rapport with the teachers, again because sometimes you share your own experiences as a former teacher. But also because of knowing the education landscape more broadly....I also feel like you've done a great job about navigating the space between the researchers and the PD providers and the teachers so you don't seem too attached to the PD team so that you are perceived as oh, you are also from UD and therefore the teachers might not be as open to being honest. (Participant D)

These quotes suggest that participants notice the evaluator's "presence" and see it as contributing to valid, meaningful data. Their own research experience likely contributed to this awareness. Participants also had suggestions for how I handled my role, discussed in the next section.

Clients also commented on my candor and curiosity. All four identified that I not only shared data (positive or negative) but also my own impressions. Even if it was not always what they wanted to hear, they seemed to value that honesty. As this story shows, sometimes delivering "the brutal truth" helped me establish trust with clients, thus facilitating use:

We had a really, really bad workshop. And before I did anything I walked over and said, this just did not go well. And you're like, no, it really stunk. I said, okay, let's talk about what's going on, and you asked really good probative questions that you don't necessarily give as answers, which was very helpful. Like what's important to you? What do you want? The first year – what's important to you? Well, then if that's important to you, what do you think you should do? Now I know. And I walk out every time saying, okay; I have an idea of what I think we should do at this point. (Participant A)

Evaluative questioning and helping clients identify or clarify their goals and make decisions based on them appear to be strengths for me. Again, these skills were most evident in longer projects with more client contact. To illustrate, contrast the depth of the story above with this comment from another client, describing the same skills but in much less depth:

I know that in general, you asked questions that helped us to think about what our content was, and whether we were getting the point across. (Participant C)

Participants valued that I asked "the right questions," and some identified that I learned quickly about their programs and content areas.

What suggestions for improvement did they offer?

Participants' constructive criticism largely mirrored the ideas discussed above, about both evaluation and evaluator. Clients in the one-year project felt constrained by the timelines and expressed some skepticism about the value of the questions the evaluation was able to ask and answer (Standard U5, Relevant information). For instance, one argued that participant satisfaction does not mean much on its own. Another commented:

I feel like the evaluation doesn't give me all the information that I need, you know, to produce a more comprehensive picture. (Participant D)

Finally, another client stated that factual data like participant demographics were not as valuable to him/her; s/he could have generated those without an evaluation team. All these critiques were framed in the context of understanding evaluation limitations and requirements.

All four clients identified data that was *not* part of the evaluation that they would have appreciated. In a sense, these comments are missed opportunities for use for either program management or research purposes. Usually these gaps were additional qualitative data (e.g., more in-depth follow-up interviews, larger number of interviews, or interviews with stakeholders that we did not include). One client felt that evaluation reports could and should include thicker description from field notes and more quotations from interviews. Survey data had inherent limitations in such small programs; it was harder to see patterns. Two clients stated that data would be more useful if it could be disaggregated (e.g., by workshop session or content area).

Regarding my professional role, several spoke about how I “grew into” it. In general, they preferred when I engaged more actively with participants and staff

members than when I took a more removed, “fly on the wall” perspective. This included my physical position during observations:

I really think like the first two cohorts, you sat in the corner. And that was partly my perception. This is what an evaluator does. And over time, as we started to perceive you as a resource, we started talking about the limits of how you could interact. (Participant A)

In this project and in one other, staff sometimes asked me to participate (e.g., model an activity, join a group to balance numbers, answer questions related to surveying) and appreciated when I did. One message in my data seems to be that it is more comfortable for project staff if the evaluator is not back “in the corner,” and that more active interaction promotes more use. However, to balance this idea, some participants also identified possible risks and boundary issues of having an evaluator too involved in project activities.

So, I think it was a very interesting role that you played, whether it was – fits into your role as a researcher or not. I think that’s an interesting part, because it could become a quasi-research and participant in the work that plays out.

You kind of set the path at some level and we set the goals, and if they don’t....it’s complicated, complicated. (Participant B)

As described above, I gradually developed relationships with clients at the same time as I was building my program evaluation skills. While no participant told me that I should have “grown faster,” one inference may be that there were some missed opportunities early in each project.

In at least one case, I could also have communicated more pro-actively to strengthen relationships and trust, and address evaluation resistance. This client did not offer direct critiques of me, rather reflecting on her role. However, his/her comments also show my areas for growth:

Should I work with an evaluation team again, I would definitely want to spend some better get to know you time, in terms of what our area of interest is and what we feel like is the purpose of the project and how we might go about doing what we do, because a little better communication about how we perceive the project and how you perceive the project, so we can work on bridging that a little earlier. [That would have] helped bring us all to the table to discuss what is going on, what is reasonable to expect, what are the best ways to measure it. (Participant B)

Discussing how different stakeholders “*perceive the project*” may be especially important where research is being conducted alongside the evaluation. These situations also require negotiation: who’s collecting what data, and what will be done with it? One client identified both redundancies and gaps in data collection, suggesting the need for better coordination. This addresses standards U2 (Attention to stakeholders) and U3 (Negotiated purposes).

In general, I could have done more to educate my clients about program evaluation – what it is, what it is for, how it can be used, and the resources it requires. I was still learning about all of this myself and not yet able to help others. Had I been more experienced, I might have been able to facilitate more process use. As one person suggested, clients might benefit from a model (e.g., a sample report or a ‘case study’ of how evaluators and project staff worked together). I also heard the need for more concrete information about evaluation timelines, and budgets.

I think that we would all have benefitted from a little better understanding of how much time could be allotted, if expectations were there. (Participant B)

Clients appreciated being “in the loop” about what the evaluators were doing, data collection schedules, response rates, etc. However, this kind of transparency has to be balanced with protecting respondent confidentiality. Some of the requests shared would have violated our agreements in collecting data (i.e., IRB).

As discussed earlier, clients identified data they wished could have been collected but was not. Sometimes this clearly exceeded the timeframe of the grant, but in other cases it might have possible to adjust plans or re-allocate resources. This suggests the potential for more stakeholder participation in the evaluation, although this would require negotiation in larger projects (i.e., different stakeholders might have competing views or desires). As described above, I found very limited evidence of process use. This also suggests a missed opportunity and a possible benefit of more participatory methods.

What are the implications of these findings for my own understanding of program evaluation and its link to improving professional learning?

In this section, I synthesize and reflect on the findings presented above.

Overall, my interviews confirmed that program evaluation can contribute to improving professional learning design, implementation, and leadership. Looking more closely at examples of use, as well as missed opportunities for it, also yields ideas for increasing the utility of program evaluations in professional learning. These lessons may be of interest to both evaluators (who want their work to be useful) and practitioners (who want to leverage evaluation for program improvement). I end with reflections on what I personally learned through conducting this study.

Evaluators can influence professional learning by drawing attention to research and best practice in the field. This can strengthen the initial design of professional learning activities and/or generate adjustments and improvements during program implementation. An evaluator's opportunity to shape the original design depends on his/her engagement early in the project (e.g., in the proposal writing phase), ongoing involvement in a longer project (e.g., where new activities emerge) or repeated

contracts with the same team (e.g., as in one of my projects, where the same investigators had multiple one-year grants). By designing data collection instruments aligned with professional learning standards and research, and sharing feedback in a candid and ongoing fashion, evaluators can help project staff “operationalize” these concepts. This can encourage instrumental use (e.g., programmatic decisions, such as designing a workshop agenda to feature more active learning) as well as conceptual use (e.g., gaining an understanding of how adults learn). Using research to anchor the evaluation aligns with Utility Standard 4 (Explicit Values) and also helps achieve relevance (Standard U5) and meaning (Standard U6).

This contribution may be particularly important where clients are not well versed in professional learning themselves. For example, some DERDC clients (including some in this study) are experts in their academic fields but new to designing professional learning activities for K-12 educators. Others may be managing a program for the first time. My data suggest that I had more influence, and my work was more used, in cases where clients knew they had a lot to learn and were eager to partner with me to learn it. On the other hand, I also found that clients who themselves were experts in professional learning could use the evaluation in other ways. For example, evaluation could provide external confirmation for a decision to make an activity align better with research (e.g., limit its scope) or could help that client educate his/her colleagues on a project. Information from the evaluation may not have been brand new to these clients, but it could still be useful.

Another role evaluation can play is giving voice to participants or other program stakeholders. This requires a sensitivity to who those stakeholders are (Utility Standard 2), the ability to engage and listen to them, and to communicate their

perspectives accurately. Findings from this process can influence program design (i.e., to better identify and plan to meet their needs) or implementation (i.e., to make adjustments based on emergent needs or participant response to professional learning). Through this study, I learned that clients most valued in-depth data directly from participants or from others familiar with their work (e.g., supervisors). In these projects, this typically meant qualitative data. Such data were seen as the most meaningful and relevant for understanding the impacts of professional learning. They were thus more useful to program leaders, who also recognized that they might not have the resources to gather such data themselves and also that participants may also feel more comfortable being candid with an external person.

Connected to Utility Standard 1 (Evaluator Credibility), it was a relief to find that I was seen as credible and competent for this type of evaluation. As an educator with a humanities background entering a social science field, I was very aware of my own learning curve. It was affirming to learn that my observation, interpersonal and communication abilities were applicable and that clients noticed my qualitative skills more than they registered any quantitative limitations. At the same time, I recognize the limitations of this analysis. My sample included only staff members who ran projects where I played a major role, and with whom I had ongoing communication. Some selection bias is inevitable.

Also related to stakeholders, research on evaluation use suggests that participatory approaches increase utility. My study yields limited insight on this issue, since all four of the evaluations I studied were relatively traditional in design. Involving program managers or other clients did increase buy-in, especially in situations where they were less familiar or comfortable with evaluation, but clients

were also candid about their own time and resource limits. Negotiating the purpose of the evaluation (Standard U3) in a general sense was important, but it is not clear that clients could have, or wanted to, devote more time to the evaluation day-to-day.

Another research consensus is that effective communication promotes evaluation use. I found this to be true, in both formal and informal ways. Standard U7 refers to “timely and appropriate communicating and responding.” I exerted great effort turning around evaluation reports quickly and making them as professional as I could. Through this study, I learned that clients appreciated this and evaluation reports were sometimes used directly, for instance, for the instrumental purpose of reporting to a funder. In other situations, a quick turnaround facilitated formative use (e.g., month-to-month program adjustments) and allowed the evaluation to be relevant (Standard U5). I also discovered some uses of evaluation reports that I had not anticipated. For example, in projects with a research component, my findings were used to bolster data collection and integrated into proposals and manuscripts. I knew this happened but did not previously classify it as a form of “evaluation use.”

On the other hand, quite often the “timely and appropriate” communication channels were not reports but conversations between clients and evaluators. It surprised me to learn how much some clients valued ongoing, even quasi real-time, communication with an evaluator, and that the evaluative resource they seemed to value most were our informal questions to help them clarify what they were doing, thinking or deciding, and why. This type of communication requires interpersonal skills and trustworthiness. Interestingly, these concepts are not made explicit in any single utility standard, though they undergird each of them. For example, I learned that my own credibility (Standard U1) was supported by my willingness to tell hard truths.

Concern for Consequences and Influence (Standard U8) requires interpersonal sensitivity and political savvy.

This type of relationship takes time to build. A very clear implication of my data is that evaluation dynamics depend on project length. I knew this instinctively but was struck by how strongly it emerged. In a short (one-year) project, an evaluation might only be able to accomplish a bare minimum, such as use for accountability. I learned of a few additional uses in short projects, but they were limited and/or occurred year-to-year (technically, beyond the timeframe of the evaluation). Again, the scope of a project and its evaluation were intertwined. Similarly, evaluators and project staff members build understanding of each other, and each other's work, over time. Longer relationships are likely to be more rewarding on both sides. And clearly longer professional learning projects are likely to feature more opportunities for evaluation use, and deeper impacts in general. However, the takeaway here is not clear because neither practitioners nor evaluators control funding parameters. We could decline to respond to one-year proposals, but this comes with an opportunity cost. Instead, evaluators may have a role to play in helping funding agencies recognize the importance of time in professional learning.

This study also leads me to think about the type of role I might want in the future. Clients told me I was most effective when I was closest to the evaluator/practitioner boundary. They preferred when I was more vocal, a "critical friend" rather than an anonymous researcher in the back of the room. This feedback confirms my self-assessment that I will work best in situations where I can develop relationships and use my interpersonal skills as well as my qualitative orientation. I

want to continue learning about and practicing different approaches to evaluation, as well as to consider internal evaluator or practitioner roles.

Finally, I learned that it is possible to solicit feedback from clients once relationships have been established. Again, acknowledging that these clients may not be “typical,” they seemed quite willing to reflect on their experiences with the evaluation. I was reminded how invested program staff members are in their work. This specific method would not be suitable for every purpose; it takes time and might only be possible after an evaluation has concluded. Still, this project made me want to develop other reciprocal feedback channels. Although I always valued client communication, I previously had concerns about asking for feedback *per se*. What if they expressed concerns I could do nothing about? Was inviting their input crossing a boundary? Would they be comfortable with the question, and candid in their feedback? This project has not eradicated all of these questions. There is some risk in inviting participation and feedback, and not every aspect of an evaluation is negotiable. Yet I now have a clearer sense of the benefits of the conversation.

References

- Alkin, M. C. (Ed.). (2004). *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks: Sage.
- Doino-Ingersoll, J.A., Haley, J., Dowell, K., & Chambliss, J. (2005). Client feedback form. Retrieved May 26, 2017 from: <https://tinyurl.com/yc5jq6fc>
- Huberman, M. (1989). Predicting conceptual effects in research utilization: Looking with both eyes. *Knowledge in Society: The International Journal of Knowledge Transfer*, 2 (3), 6-24.
- Johnson, K., Greenseid, L.O., Toal, S.A., King, J.A., Lawrenz, F. & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, 30 (3) 377 – 410.

- Learning Forward (2011). Standards for professional learning. Retrieved August 6, 2017 from <https://learningforward.org/standards-for-professional-learning>
- Patton, M.Q. (2004). The roots of utilization-focused evaluation. In M.C. Alkin (Ed.), *Evaluation Roots*, 276-292. Thousand Oaks, CA: SAGE.
- Patton, M.Q. (2008). *Utilization-focused evaluation*, 4th edition. Thousand Oaks, CA: Sage.
- Patton, M.Q. (2013). Utilization-focused evaluation checklist. The Evaluation Center, Western Michigan University. Retrieved May 26, 2017 from: https://wmich.edu/sites/default/files/attachments/u350/2014/UFE_checklist_2013.pdf
- SocioCultural Research Consultants, LLC (2016). Dedoose Version 7.0.23, web application for managing, analyzing, and presenting qualitative and mixed method research data.
- Stufflebeam, D.L. (1999). Program evaluations meta-evaluation checklist. The Evaluation Center, Western Michigan University. Retrieved March 16, 2017 from: https://www.wmich.edu/sites/default/files/attachments/u350/2014/program_metaeval_short.pdf
- Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D.B. & Balcazar, F.E. (2013). Understanding and measuring evaluation capacity: A model and instrument validation study. *American Journal of Evaluation*, 34 (2), 190 – 206.
- Weiss, C.H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19 (1), 21 – 33.
- Yarbrough, D.B., Shulha, L.M., Hopson, R.K. & Caruthers, F. A. (2010) *The Program Evaluation Standards*, 3rd Edition. Thousand Oaks, CA: SAGE.

Appendix 1

Analysis of evaluation use by PD programs Interview Protocol

Reinforce purpose of project, confidentiality of responses, and request for candor.

Questions	Dimensions
1. [Prior to the interview, participants will be sent (1) a summary fact sheet of the program evaluation, including activities I conducted and interactions I had with the program and (2) the most recent evaluation report I submitted for the program. I will bring copies of both documents and show them at the start]. Have you had a chance to review these documents? Would you like a few more minutes to do so now? [If necessary, delay start of interview while participant reviews].	Preparation
2. I know it has been a little while since we worked together. How familiar were you with the materials you just reviewed? Was any information on the fact sheet new to you? Do you remember receiving that report? Reading it?	Preparation Also touches communication clarity, stakeholder involvement
3. Let's set the context. Why was DERDC conducting an evaluation of [program name]? (Probe on reason for an evaluation in the first place as well as decision to hire DERDC) How much experience did you/the program have with evaluation?	Political climate
4. What did you see as the purpose of the evaluation? Was it met?	Political climate
5. How much investment did the program have in this evaluation? Who, if anyone, cared about it, and why? (Introduce idea of stakeholders: program clients, staff, managers, funders, regulators, etc.)	Decision characteristics
6. Let's start with the evaluation questions. [Review]. Have you seen these questions before? How much involvement, if any, did you have with developing them? Were these the most important things to learn about the program, from your perspective? Why/why not?	"Personal factor"
7. Now let's look at how the evaluation collected and analyzed data to answer those questions. [Review]. How were you involved in evaluation procedures? (e.g., provide information, help recruit participants, review instruments, analyze data?)	Stakeholder involvement Relevant information

15. Was evaluation communication timely? Did you have what you needed, when you needed it? Explain. (Includes formal deadlines, but also information provided in time to guide decisions)	Communication clarity Meaningful products
16. DERDC has several staff members, but I was the lead evaluator for this project [/part of project]. Help me think through my credibility as an evaluator. What relevant skills or experiences did you think I brought to this evaluation? What could have made me more credible?	Communication clarity Timely and appropriate communicating and responding
17. If not already covered, pose a question about other interpersonal and cultural competence skills.	Timely and appropriate communicating and responding
18. What else should I have asked you that I did not? Please share any other feedback or comments about the evaluation of this program.	Evaluator credibility
	Evaluator competence
	Other/closure

Appendix L

PROPOSAL NARRATIVE

Looking at Teacher Learning from the University: Levers for Improvement Overview

My Educational Leadership Portfolio (ELP) will address and unpack a multidimensional question: how can universities best facilitate professional learning for educators in the K-12 public system? This proposal outlines the **need** to strengthen professional learning, the **problems** currently associated with designing, implementing, and evaluating professional learning opportunities and the **resources or levers** that universities can bring to bear in this essential work. I focus on the Delaware public education and higher education **organizational context**. Reflecting my varied professional background, I investigate the intersections of professional learning or development⁴, program evaluation, and educational leadership. I hope to show how universities understanding professional learning can yield more robust programs or initiatives, with more effective implementation and leadership. My artifacts, detailed later in this proposal, represent a variety of strategies to connect research to practice and university resources to K-12 education, with the purpose of improving professional learning and in turn instruction and student achievement.

⁴ A more extensive definition of these terms occurs later in this proposal. Note that both research and practitioner communities are increasingly choosing “professional learning” instead of “professional development” (PD). According to the Delaware Department of Education (2016), this shift is intended “to connote the importance of continuous improvement.” Where possible, I follow this vocabulary although “PD” also appears in many places due to its prevalence in past literature.

Organizational Role

I currently work part-time for two centers at the University of Delaware (UD) whose missions are to connect the resources of the university with the needs of the education sector. In this project, I focus on the K-12 public education community, particularly in the state of Delaware. These centers are the Delaware Academy for School Leadership (DASL) and the Delaware Education Research and Development Center (DERDC), both housed within UD's College of Education and Human Development (CEHD). DASL provides research-based professional and leadership development programs and services. DERDC conducts program evaluations and applied research related to education; my work there focuses on professional learning initiatives. I am also the preceptor for two classes in the Master's in Teacher Leadership Program, also within CEHD. Through my various projects at UD as well as my decade of varied prior experience in the field of education, I understand professional learning through multiple roles: participant, designer, facilitator, manager, and evaluator.

In this ELP, I plan to investigate how those of us outside K-12 schools and districts can add value to the professional learning that occurs within them. How can we translate insights from research to inform stronger programs or practices, helping to achieve better teaching and learning outcomes? How can we ensure that research or policy efforts are grounded in the real world of schools? Conversely, how can a busy district, school or teacher leader integrate a more research-based perspective into her responsibilities to plan, deliver, and evaluate professional learning? My proposed artifacts, discussed below, are efforts towards understanding these questions and tools for bridging the research to practice divide.

Organizational Context

University Context

This ELP focuses primarily on Delaware K-12 public education, but as I work in a university I begin with an overview of that context. UD engages with the broader community, including K-12 education, in many ways through each of its seven colleges. In 2015, the university was awarded the Carnegie Community Engagement Classification in recognition of these efforts. In its Carnegie application, UD demonstrated strong investments in community engagement. For instance, a campus-wide employee survey showed that faculty committed 26% of their work time to this purpose; for professional staff, it was 24% (University of Delaware, 2014). The university also laid out a rationale for community engagement:

UD recognizes that community engagement is critical not only to our public service mission but also to our educational and research missions: enriching student learning, improving the effectiveness of our teaching, allowing partnerships to guide research, and enhancing the impact of scholarship (University of Delaware, 2014, p. 2).

Using this broad framework, the university's partnerships with K-12 public education are seen as mutually beneficial. The ultimate aim of improving student learning outcomes in Delaware clearly supports UD's "public service...educational and research missions." More proximally, by getting involved with K-12 professional learning through strategies such as those discussed in this ELP, university personnel not only serve the public but also build relationships and gain insight into educational contexts, which can yield research partnerships, opportunities to develop new educational interventions or research/evaluation methods, audiences for dissemination

and application of research, networking for their own development, etc. Thus, university and school, district or state capacities are reciprocally strengthened.

There is increasing interest and momentum for such partnerships in Delaware. In 2015 UD unveiled a new strategic plan, which included community engagement as a key initiative and made the following priority recommendation (emphases mine):

Engage more Delaware pre-K–12 teachers and their classrooms in UD research/outreach programs, the traditional and performing arts, **teacher professional development** and innovative curricula (University of Delaware, 2015).

To strengthen and organize existing partnerships between higher and K-12 education as well as establish new ones, in 2016, UD launched the Partnership for Public Education (PPE). While PPE is a new university-wide coordinating and mobilizing structure, UD also has numerous research and public service centers that serve as connection points between the university and the broader community. Both of the centers for which I work, DASL and DERDC, serve this function within the College of Education and Human Development (CEHD). According to its mission statement, CEHD

develops solutions to the problems that confront our schools and the challenges encountered by our children, youth, and families. Although our primary mission is to conduct research and train UD students to become highly qualified professionals and leaders in their fields, we also partner with organizations and agencies to ensure that Delaware children, teachers and families receive the best possible education and vital social services (2013).

In 2013, the College estimated that its faculty and staff provided professional development or training to 10,000 early childcare providers, 2,000 educators and 500 school leaders in Delaware and neighboring states. In summary, my ELP's focus on

professional development collaborations between higher education and K-12 is consistent with the organizational structure and strategic direction of CEHD and UD.

Public Education Context

I now examine public education in Delaware, situate it within the national context, and begin to build connections to professional learning needs. Delaware has 226 public schools, together enrolling 136,027 students and employing 9,064 teachers and 835 administrators (Delaware Department of Education, 2016b). Teacher mobility is lower in Delaware than in the country as a whole, but there is still substantial turnover. In the 2015 hiring season, 964 teachers (approximately 11%) were new to their schools (Robertson-Kraft & Hejlek, 2016). Statewide, 23% of teachers have fewer than five years of experience (Rodel Foundation , 2015). Turnover increases the need for teacher development.

Delaware education is characterized by standards, assessments, and educator accountability, all of which also have implications for professional learning. The state adopted the Common Core State Standards for Math and English Language Arts and the Next Generation Science Standards, and all three new sets of standards were expected to be fully implemented in 2016-17. Assessments have changed rapidly as well. The Smarter Balanced assessment is administered for grades 3-11, and a new, NGSS-aligned science assessment is in development. Educators are evaluated through a statewide system (DPAS II), which includes five components, one based on student achievement outcomes. Delaware recently finished implementing a multi-year Race to the Top (RTTT) federal grant, which supported these new standards, assessments, and accountability systems. Delaware also invested RTTT funds heavily in “human capital” initiatives related to teacher and leader recruitment, development, and

evaluation. Although performance improved for some students, gaps persist. At a recent statewide education conference, a consensus emerged that “the past few years’ education reforms have not done enough to help disadvantaged children catch up, so Delaware’s education system should make closing the achievement gap a top priority” (Albright, 2016). Doing so requires enhanced educator practices.

Delaware mirrors national trends: rapid change, increasing standards, gaps and inequities, and complex learning needs for both students and teachers. Technological innovations open new possibilities and demand that instruction evolve to meet the needs of 21st century learners (Jacobs, 2012). New knowledge about teaching and learning is constantly being produced. Fueled by concerns about global competition, our nation’s academic aspirations are increasing requiring profound shifts in practice. We are still adjusting to new standards and assessments, while the implications of new legislation such as the Every Student Succeeds Act (ESSA) are as yet unknown. The human context is dynamic, too. America’s public school students are increasingly diverse in culture and language; growing numbers live in poverty (Layton, 2015a). Harnessing student diversity as an educative resource while confronting the persistent inequities and achievement gaps in the educational system is a complex challenge. The educator workforce continues to churn; 16% of teachers either leave the profession or change schools annually (Goldring, Taie, Riddles & Owens, 2014). Faced with such turbulence and such high stakes, the need for learning on every level – classroom, school, and system – is acute.

Standards and Definitions of Professional Learning

In response, those designing educational programs almost inevitably rely on activities to build educators' knowledge and skills. As Desimone (2009) puts it, "education reform is often synonymous with teachers' professional development" (p. 181). This is true both nationally and in the state. In 2016, the Delaware Department of Education (DDOE) launched a grant competition for "Reimagining Professional Learning," which reflected the agency's awareness that "In the past three years, both state and local level approaches to professional learning have evolved in many meaningful ways, and yet there is still so much to do to realize the commitment to reimagined, top-notch professional learning for every Delaware educator as the norm" (Delaware Department of Education, 2016a). In addition to recognizing the urgency of professional learning, DDOE also put forth this definition:

Professional learning is defined as a comprehensive, sustained, and intensive approach to improving teachers' and principals' effectiveness in raising student achievement. There is a shift from the concept of professional development to professional learning to connote the importance of continuous improvement.

Around the same time, the federal government issued an updated definition of professional development in the context of the Every Student Succeeds Act (ESSA).

The term 'professional development' means activities that—

"(A) are an integral part of school and local educational agency strategies for providing educators (including teachers, principals, other school leaders, specialized instructional support personnel, paraprofessionals, and, as applicable, early childhood educators) with the knowledge and skills necessary to enable students to succeed in a well-rounded education and to meet the challenging State academic standards; and

"(B) are sustained (not stand-alone, 1-day, or short term workshops), intensive, collaborative, job-embedded, data-driven, and classroom-focused" (Learning Forward, 2015)

The law goes on to enumerate three main purposes for PD: to build educator understanding of academic content, student learning, and methods for assessing that learning and adjusting instruction. It requires that PD be “evidence based” and “personalized.” The DDOE and ESSA definitions both show that PD should not be viewed as a discrete event, but as a complex process of adult learning, with the purpose of promoting student achievement. These definitions align with standards for professional learning.⁵ In 2012, the Delaware Professional Standards Board adopted the Learning Forward Standards (2011). These are shown in Figure L1, below.

Delaware’s Standards for Professional Learning:	Professional learning that increases educator effectiveness and results for all students...
Standard: Learning Communities	Occurs within learning communities committed to continuous improvement, collective responsibility, and goal alignment.
Standard: Leadership	Requires skillful leaders who develop capacity, advocate, and create support systems for professional learning.
Standard: Resources	Requires prioritizing, monitoring, and coordinating resources for educator learning.
Standard: Data	Uses a variety of data sources and types of student, educator, and system data to plan, assess, and evaluate professional learning.
Standard: Learning Designs	Integrates theories, research, and models of human learning to achieve its intended outcomes.
Standard: Implementation	Applies research on change and sustains support for implementation of professional learning for long-term change.
Standard: Outcomes	Aligns its outcomes with educator performance and student curriculum standards.

Figure L1 Learning Forward and Delaware Standards for Professional Learning

⁵ Indeed, Learning Forward strongly influenced the development of the ESSA definition (Hirsch, 2015).

These standards and definitions provide the context for my ELP. As I will show in the next section, they are derived from a large body of research about professional learning. While this research demonstrates the potential of professional learning to improve instruction and student achievement, it also shows that too often it falls short of these goals.

Problem Statement

As described above, professional development for educators is ubiquitous as an improvement strategy. Examples of successful reforms highlight the critical role of PD. As Guskey (2000) states, “one constant finding in the research literature is that notable improvements in education almost never take place *in the absence of* professional development” (p. 4). Professional development is also a huge industry, as state, district and school leaders and other decision-makers chase its promise. In 2012, Education Secretary Arne Duncan estimated the federal PD investment at \$2.5 billion per year (Layton, 2015b). In a recent study, TNTP found that three large districts poured nearly \$18,000 *per teacher per year* into professional learning (TNTP, 2015). Unfortunately, the return on all this investment is paltry and inconsistent. PD is far from a miracle cure or magic bullet. That is the conclusion of numerous research syntheses and policy studies (Borko, 2004; TNTP, 2015). Teachers concur – most give their professional development experiences mixed if not downright scathing reviews (Calvert, 2016; TNTP, 2015).

The promise of professional *learning* spurring more effective instructional practices and increased student achievement often seems like, as TNTP titled its report, a “mirage.” In this ELP I join other researchers and practitioners in efforts to solidify that haze. I look at reasons why professional learning is so often insufficient

and what can be done about that, especially by those of us working in universities yet primarily concerned with what happens on the ground in K-12 schools. My career in education has spanned sectors and taken me from classroom teaching to program management and most recently program evaluation and educational research. Because of these varied experiences, I am oriented towards university contributions that make a real difference on the ground for teachers and students. Through my experiences as a participant, developer, facilitator, and evaluator of professional learning initiatives, I have also become convinced that research suggests powerful ways to improve them. As sites of engaged and applied research, universities like UD can use several levers to effect positive change in professional learning. Later in this proposal, I identify and give examples of five. In various ways these all draw on the large and constantly developing body of research on the subject, and are aligned with state and national standards. The literature on professional learning is vast, and a full review is beyond my scope here. Subsequent artifacts look more deeply into what we know, and still need to learn, about professional learning.

Research on Professional Learning

Since about 2000 there has been an effort to study the effectiveness of professional development on student learning through large, rigorous, sometimes experimental studies, and to look more specifically at what design features of PD make the difference. Recent meta-analyses (Blank and de la Alas, 2009; Yoon, Duncan, Lee, Scarloss & Sharpley, 2007), research syntheses (Desimone, 2009; Borko, 2004) and large studies (Garet, Porter, Desimone, Birman & Yoon, 2001; Penuel, Fishman, Yamaguchi & Gallagher, 2007) have proposed the following features of effective PD:

Content focus – Effective PD involves teachers in reinforcing academic concepts, then learning how to convey those concepts to students. So for example, in a math PD, teachers would do math and discuss pedagogical strategies for specific math concepts.

Time/duration and context – Research is clear that teacher learning takes time. For example, of 16 programs identified that demonstrated significant gains in student achievement (Blank & de las Alas, 2009), the average total time was 91 hours and the average duration was 6 months. Some programs had more than 100 hours and lasted 16 months. Related, professional learning is more powerful when embedded within the real context where teachers work.

Active learning – Effective PD engages teachers in a variety of activities that apply concepts to their practice. This could include hands-on activities or lesson planning time during workshops, or less traditional PD experiences such as observing fellow teachers, being observed/coached, participating in school improvement or curriculum development processes. “Sit and get” workshops are the antithesis of effective PD.

Coherence – Effective PD surfaces, and attempts to connect with, what teachers believe and know already (Desimone, 2009; National Research Council, 2000). It also aligns with the existing structures within which they work, such as state/national standards, curricula, and other concurrent reform initiatives.

Collective participation – Involving teams of educators from the same system (school, district) has shown to be an effective design. It increases collaboration and peer support, makes it more likely that changes in instruction will “stick,” and creates momentum for change.

These research-based features are practical and can be used to inform PD design and delivery. Yet emerging research also cautions us not to apply them simplistically. In a review of 28 experimental studies, Kennedy (2016) argues that we need to look beyond the mere presence or absence of PD features. For example, conventional wisdom suggests that job-embedded supports like coaching and collaborative structures such as professional learning communities (PLCs) would yield strong PD. Kennedy raises common-sense, yet easily overlooked, questions about quality and context. For example, how well trained are the coaches? What happens during coaching? Are PLCs engaged in meaningful or contrived work? What is the level of “buy in”? She concludes:

We need to replace our current conception of “good” PD as comprising a collection of particular design features with a conception that is based on more nuanced understanding of what teachers do, what motivates them, and how they learn and grow. We also need to reconceptualize teachers as people with their own motivations and interests. The differences shown here among PD methods of facilitating enactment strongly suggest the importance of intellectually engaging teachers with PD content, rather than simply presenting prescriptions or presenting bodies of knowledge (Kennedy, 2016, p. 30)

Similarly, Timperley (2011) argues that more important than any particular design is teachers’ level of engagement in professional learning and what they do, or do not, learn from the experience. The ideas of teacher agency and engagement, as well as models of teacher learning, will be explored further later in this ELP.

Unfortunately, many of the opportunities available to teachers fall far short of these research recommendations, and do not fit the standards and definitions presented above. Brief, even “one shot” sessions are common; workshops are the most common delivery format (Gates Foundation, 2014). Echoing Kennedy’s argument, many teachers report dissatisfaction with PLCs and lesson observations, suggesting that what

could be high-impact professional learning opportunities are not being implemented effectively (Gates Foundation, 2014). Professional development is often “one size fits all,” with teachers having few opportunities to make choices about what they learn or to differentiate or personalize the activities to their specific teaching assignment, level of experience, or student needs (Calvert, 2016; Gates Foundation, 2014). In general, too many educators express that PD is not preparing them to address critical problems of teaching and learning. These data are derived from large national surveys of educators.

An analogous state survey provides insight into educators’ perceptions in Delaware (New Teacher Center, 2013a). The Delaware Teaching, Empowering, Learning and Leading (TELL) was administered to licensed educators in districts and charters statewide. The survey was online, anonymous with a response rate of 59%. Results related to PD are shown in Figure L2, below.

Q8.1 Please rate how strongly you agree or disagree with statements about professional development in your school.	
a. Sufficient resources are available for professional development in my school.	67.8%
b. An appropriate amount of time is provided for professional development.	66.9%
c. Professional development offerings are data driven.	76.0%
d. Professional learning opportunities are aligned with the school's improvement plan.	82.9%
e. Professional development is differentiated to meet the needs of individual teachers.	44.3%
f. Professional development deepens teachers' content knowledge.	60.2%
g. Teachers are encouraged to reflect on their own practice.	86.1%
h. In this school, follow up is provided from professional development.	55.8%
i. Professional development provides ongoing opportunities for teachers to work with colleagues to refine teaching practices.	64.5%
j. Professional development is evaluated and results are communicated to teachers.	42.4%
k. Professional development enhances teachers' ability to implement instructional strategies that meet diverse student learning needs.	71.2%
l. Professional development enhances teachers' abilities to improve student learning.	76.2%

Figure L2 TELL Delaware data

While many of these responses seem relatively affirming, there are areas for concern. Differentiation of PD to individual needs and follow-up from PD both have relatively low levels of endorsement. Furthermore, disaggregated analyses indicated that respondent groups viewed professional development differently, and that teachers were more critical. For example, fully 98% of administrators agreed that professional development provides ongoing opportunities for collaboration but only 62% of teachers agreed. Ninety percent of administrators stated that follow up from professional development is provided, compared to just over half of teachers (54%) (New Teacher Center, 2013b). Why do administrators believe design features are in place but teachers do not experience them? These disparate data suggest the value of exploring stakeholder perspectives on professional learning more deeply and attending the complexities of implementation. They may also indicate the need to involve administrators more fully in professional development.

Evaluation of Professional Learning

The TELL data also indicate that evaluation of professional learning is a weakness. Just 42.2% of respondents agreed that “professional development is evaluated and the results are communicated to teachers.” There is room for growth in the design and delivery of professional learning opportunities – and *relatedly* in their evaluation. A central assumption of my ELP is that more rigorous, meaningful data about professional learning can be leveraged into improvements that benefit teachers and, in turn, students. As the leading scholar in PD evaluation, Guskey (2000), puts it:

A lot of good things are done in the name of professional development. But so are a lot of rotten things. What educators haven't done is provide evidence to document the difference between the two. Evaluation provides the key to making that distinction. (p. 94)

In this next section, I review frameworks for the evaluation of professional learning, and outline some ways to enhance current evaluation practice. Later artifacts build on these ideas.

There is a general call for more rigor in research about PD. Of 1300 studies reviewed by Yoon and colleagues (2007), only nine met *What Works Clearinghouse* evidence standards. A chorus of researchers states the need for better-designed and more rigorous studies of PD (Borko, 2004; Desimone, 2009), with specific attention to impacts on student achievement. This push is also reflected in the evaluation criteria for federal and state grants and funding. For example, the DDOE application for the Reimagining Professional Learning Innovation Grant states that proposals must specify “systems for gathering and analyzing evidence of impact of professional learning on teacher practice AND student learning outcomes” (Delaware Department of Education, 2016a).

Current practice in evaluating PD programs often falls far short of these expectations. Most commonly, PD is evaluated by surveying participants to see if they enjoyed the experience. Participant satisfaction is important and necessary but insufficient for understanding the deeper impacts of PD. Reeves (2010) frames the problem starkly: “The central challenge for educational systems around the world is the substitution of effectiveness for popularity” (p. 2). We need to dig deeper in evaluation, and Guskey (2000) provides a model for how to do so. According to Guskey, evaluation of PD should attend to five levels of effects: participants’ reactions, participants’ learning, organization support and change, participants’ use of new knowledge and skills, and student learning outcomes. Each level is a prerequisite

for the others yet success at one level does not necessarily mean success at the next. For example, a participant can love a PD experience and learn a lot in it (Levels 1 and 2), but if his/her context fails to support the new learning (Level 3), s/he may never demonstrate changes in teaching practice (Level 4) and student outcomes may not improve (Level 5). This could happen if structures such as the school schedule, teaching/PLC arrangements, or curriculum requirements compromise the implementation of PD, or if school leaders or fellow teachers did not understand or support it. (This is an issue of coherence). To understand the intricacies of PD transfer, it is necessary to evaluate all five levels. As the levels increase, evaluation becomes more complicated and time-consuming and yet the information gathered is more valuable. Guskey argues that even organizers of small PD programs can inquire into all five levels, and external evaluators certainly should. The Guskey framework for evaluating PD is the most prominent, and indeed DDOE endorses its use (Delaware Department of Education, 2016a). However, other models do exist and some of them explicitly build upon or complicate Guskey, usually by presenting teacher learning in cyclical or iterative rather than linear ways (Timperley, 2011; Coldwell & Simpkins, 2011; King, 2014). These models are discussed in one of my artifacts (see below).

Evaluation frameworks present a systematic way to understand the intended effects of professional learning. I want to unpack some implications of using the Guskey model that are relevant for my ELP. Evaluating level 3 (“organization support and change”) involves looking at school or other organizational conditions and culture, and at how these influence the implementation of professional development activities. Here Guskey’s framework fits well with Kennedy’s (2016) argument about the importance of context. Educational leaders and other practitioners may need help or an

outside perspective to be able to accurately perceive and evaluate this context. Second, completing a full Guskey-aligned evaluation requires a substantial timeframe and budget. Teacher learning is slow and messy, and longer-term studies tend to show greater changes in instructional practice and student learning (Kennedy, 2016) yet funders may impose unrealistic evaluation timelines. Finally, although the Guskey framework is user-friendly, practitioners may need support to translate it into practice. They also may require additional resources or capacities to fully conduct an evaluation. All of these implications are also opportunities for universities to contribute to the evaluation process.

Improvement Goal

The goal of this ELP is to use research and evaluation to improve professional learning in K-12 public schools in Delaware, thus supporting more effective instructional practices and higher student achievement. I do this by identifying five levers that university-based staff members such as I can use:

1. Synthesize research related to professional learning; develop related recommendations.
2. Conduct research and evaluation related to professional learning.
3. Educate others about research-based professional learning design, implementation, and evaluation.
4. Innovate with evaluation methods and help educators develop their own evaluation capacity.
5. Increase understanding of what educators value in professional learning efforts and evaluations.

For each lever, I investigate several key questions and present one to four artifacts that either demonstrate how I have used the lever in the past or present a plan to do so in the future.

Table 1 below aligns the improvement levers discussed above with artifacts I propose for my ELP. **Bold artifacts** are submitted along with this proposal narrative.

Table L1 Proposed artifacts

Lever/Strategy	Key Questions	Artifacts
1. Synthesize research related to professional learning; develop related recommendations.	What does the latest research tell us about PD design, implementation, and evaluation? What do we still need to learn? How can practitioners use this information?	a. Review of frameworks for professional learning. b. White paper for DDOE with recommendations for professional learning for teacher leaders, based on results from the TLI study. c. PowerPoint presentation about how organizational context and school leadership influence PD implementation.
2. Conduct research and evaluation related to professional learning.	Do professional development initiatives operate as planned and achieve intended results? How can we deepen our understanding of professional learning impacts?	d. Study of the Schools That Lead Teacher Leadership Initiative e. Synthesis of evaluations of three PD programs I conducted for DERDC
3. Innovate with evaluation methods.	What are different methods to evaluate professional learning? What are tools to make these methods accessible	f. Meta-evaluation of the Specific and Innovative Improvement Practices (SIIP) Grant

	in the field?	
4. Educate others about research-based professional learning design, implementation, and evaluation.	How can we help those who are, or will be, in positions to make decisions about professional learning to understand research-based practices?	g. Reflections from teaching EDUC 774, Designing Professional Development h. Understanding by Design curriculum unit to teach graduate research assistants about PD evaluation i. Online PD toolkit (Google Sites)
5. Increase understanding of what educators value in professional learning efforts and evaluations.	What do teachers and administrators value in professional learning (and evaluation of it)?	j. Analysis of evaluation use/usefulness, focused on same three PD evaluations as artifact d. k. Content analysis of student professional development plans from EDUC 772

Artifact Descriptions

The following are overviews of each lever and descriptions of each artifact.

1. Synthesize research related to professional learning; develop related recommendations.

Artifact a: Review of frameworks for professional learning. This literature review will examine models of how teachers learn and how teacher learning is connected to changes in instruction and in student achievement. Thus, it sets a conceptual foundation for the rest of my ELP. For example, Guskey (2002) lays out a framework for teacher learning that corresponds to his evaluation framework for professional development (2000). The National Research Council (2000) contributes more of a learning sciences perspective. As I read further this fall, I discovered other

models for conceptualizing the processes of professional learning (and thus, for evaluating or measuring that learning). Some have come out in the past couple of years (e.g., Kennedy, 2016; Kennedy, 2014; King, 2014) while others were new to me (e.g., Opfer & Pedder, 2011; Kennedy, 2005; Clark & Hollingsworth, 2002). Others focus on international contexts but generate insights that made me think about US education in new ways (e.g., Timperley, 2011; Coldwell & Simkins, 2011).

Artifact b: White paper and recommendations related to teacher leadership. This is a companion to artifact c, which will be a full report of a qualitative study of the Schools that Lead Teacher Leadership Initiative (TLI). In order to make its findings more accessible, I will prepare a short white paper focusing on actionable recommendations for policy and practice. These recommendations may pertain to job-embedded professional learning such as that found in the TLI; to teacher leadership; and/or to organizational and leadership supports necessary for teacher leadership. This white paper will be shared with the DDOE, which is funding the TLI study. This artifact will demonstrate strategies for distilling and communicating study results.

Artifact c: PowerPoint presentation about how organizational context and school leadership influence PD implementation. This artifact is my final project for EDUC 890. We were asked to read and annotate five empirical articles on a leadership-related topic of our choice, then to synthesize and present what research tells us, and what still needs to be learned, about that topic. I chose to examine how organizational context, including school leadership, influence the implementation of PD. My program evaluation experiences have made me keenly interested in the ways that context facilitates or constrains teachers in using new practices. (For example, I

regularly found that participants in the same PD session had very different experiences once they went back to their schools and tried to use what they had learned). Since the course was centered on leadership, I focused on what principals/teacher leaders could do to promote implementation of PD. Students in 890 included many K-12 leaders: teacher leaders, principals/APs, even a superintendent. In addition to synthesizing research, this presentation can be seen as an attempt to educate practitioners about an important aspect of PD (i.e., it could also fit under lever 4). My presentation received mixed feedback from my classmates. They generally found the topic interesting and germane but the reasoning too abstract. One stated, in effect, “it’s hard for grasp for people who have not thought about this as much as you have.” I plan to incorporate this feedback along with any feedback from my committee when I revise.

2. Conduct research and evaluation related to professional learning.

In this ELP, I define “research” as the use of systematic approaches to answer questions and generate knowledge about professional development and “evaluation” as a subcategory of applied research which is specifically conducted on behalf of a client, to address questions about the implementation (i.e., process) or outcomes of a program in order to guide improvements. There are different schools of thought about the ultimate purpose of evaluation, but I strongly embrace the “use” paradigm. Sometimes external outcome and/or process evaluations are required by the terms of a government or foundation grant. This is the case for most of my DERDC evaluations, examples of which appear in my ELP. Other times programs may choose to commission evaluation to answer specific questions or to document their impacts. In either case, the university contributes human, technical and intellectual resources for evaluation.

Artifact d: Study of the Schools That Lead Teacher Leadership Initiative.

I am currently conducting a study of the Teacher Leadership Initiative (TLI), a program of Schools That Lead (STL), a nonprofit organization in Delaware. The TLI brings together cohorts of teachers for an intensive professional learning experience spanning two years and structured as an inquiry project. Participating teachers select a Student Learning Question, and then work with peers to collect data, reflect, and establish next steps to improve instruction. Over time, TLI participants are also expected to scale the work and this process to several other teachers.

STL approached DERDC in 2016 about the feasibility of a research partnership. STL wanted to learn from the experiences of its first TLI participants and to document the early impacts of this program. The DERDC Director decided to allocate some of the center's state line funding to support a study. The study will focus on the first cohort of TLI schools and participants and address the following research questions:

1. How do TLI participants describe their experiences within the program? To what extent are these experiences consistent with the Shared Learning Framework?
2. How do relationships between TLI participants, their peers, and their administrators develop during the TLI?
3. How does participating in the TLI influence the instructional practices of participants? Others?
4. What impacts are evident on student learning?
5. What organizational conditions and supports (e.g., trust, leadership) facilitate or constrain the work of TLI participants?

The purpose is to qualitatively study the experiences and document the impacts of the TLI on the first two cohorts of participants, and to generate recommendations to

inform future research and other initiatives related to teacher leadership and professional learning in Delaware. The study will use primarily qualitative methods, including interviews of TLI completers (i.e., members of Cohort 1), teachers to whom they “branched” in the program’s scaling model, and administrators and other leaders within their schools. The intent is to learn from a range of contexts (e.g., grade levels or other teaching assignments, school demographics). Although it is expected that interviews will form the bulk of the data in this study, other data sources (e.g., artifacts or documentation produced by participants in the course of their TLI experience) may be reviewed.

Artifact e: Synthesis of evaluations of three professional development projects. The purpose of this artifact is to demonstrate how I have used this lever, i.e. conducted evaluations of professional development. Originally, I planned to simply submit a completed evaluation report of a PD program. However, with encouragement from committee members I decided to do something more ambitious that takes advantage of my breadth of experience and gives me a chance to reflect and learn. Between 2012 and 2015 I evaluated three professional development programs on behalf of DERDC:

MADE CLEAR - a large, NSF-funded project aiming to embed climate change education into the formal K-12, informal, and higher education sectors in both Delaware and Maryland. For consistency with the rest of my ELP, I will focus on the project’s PD for in-service K-12 teachers;

Delaware Transition to Teaching Partnership (DT3P) - funded through a multi-year grant from the USDOE, DT3P is an alternative route to certification program for teachers of ELA, math, or science in high-needs secondary schools. This project

combines in-person intensive professional development and graduate courses with coaching and other job-embedded supports.

Title II - this is an annual one-year grant from the DDOE to researchers and professional developers at UD. I evaluated it for three consecutive years (2013, 2014, 2015). While the specifics varied from year to year, Title II always included a menu of different content-based PD opportunities helping teachers adjust to new standards and assessments.

I plan to look across my evaluations of MADE CLEAR, DT3P and Title II and to distill insights related both to PD design and PD evaluation. I will need to develop specific questions and methods for the analysis/synthesis. These might include:

PD design

- To what extent do the programs demonstrate the characteristics of effective PD? e.g. content focus, active learning, collective participation, duration/length/context, coherence.
- What lessons can we draw about participants' responses to and learning during in-person PD? (i.e., Guskey Level 1 & 2)
- All three programs included intensive one- or multi-day PD institutes.
- What lessons can we draw about PD implementation/follow up? (i.e., Guskey Level 3)
- What supports facilitated teachers' efforts to use the knowledge/skills acquired from the PD, and what barriers did they face?
- What lessons can we draw about participants' use of new skills/knowledge? (i.e., Guskey Level 4)

PD evaluation

- What evaluation questions and methods did I use most frequently?
Which were particularly informative?
- What variation exists in the evaluation questions and methods? Why did some projects permit questions/methods that others did not?
- Specifically to what extent was I able to evaluate the higher levels of the Guskey framework (i.e., changes in instructional practice and student learning)?

This artifact will connect to artifact k, in which I obtain feedback on my evaluation efforts from the managers of these three PD programs.

3. Innovate with evaluation methods.

Another contribution universities can make is to develop and disseminate new ideas. In my field, this entails developing innovative methods for evaluating professional development programs. I have limited experience using this lever so far but I include one artifact that, I hope, points the path forward.

Artifact f: Meta-evaluation of the Specific and Innovative Improvement Practices (SIIP) Grant In 2013-14, the DDOE funded fourteen projects under a new competition, the Specific and Innovative Improvement Practices (SIIP) grant. These innovative and promising projects focused on one or more of the following goal areas: (1) teacher-led projects that drive improved student outcomes (2) Common Core implementation and assessment (3) student supports and dramatically improved school climate, and/or (4) accelerating the achievement of underperforming groups. To reach these goals, projects almost all included one or more professional development activities, thus SIIP fits well within the focus of my ELP.

DERDC conducted an evaluation of the SIIP grant. In order to synthesize across 14 different projects, we developed evaluation rubrics for implementation and outcomes of the SIIP projects. In essence, we were conducting meta-evaluations, using the 14 SIIP project reports as our data source. To my understanding, this was the first time this procedure had been used in the state. This was a collaborative project. The DERDC Director led this project and initially conceptualized the method. Together with a graduate research assistant, I drafted the evaluation rubrics. As a team of three we then analyzed and interpreted data together and co-authored the report. We received positive feedback on this evaluation. Another aspect of this project was providing technical evaluation assistance to SIIP project managers and to the state. We met once with each project early in the grant year and several times with DDOE and made ourselves available for further support. We aimed to help local projects and the state develop their own evaluation capacity. This goal was only partially met.

4. Educate others about research-based professional learning design, implementation, and evaluation.

This lever involves using a university's teaching mission to improve professional learning. I have used this lever both formally and informally as an instructor in the Masters in Teacher leadership program, as a mentor for graduate research assistants, and as a student myself. In the first situation, I contributed to a new course about professional development. In the other two, I created new opportunities to engage others in learning about professional development.

Artifact g: Reflections on teaching EDUC 774. I was a preceptor for EDUC 774, *Designing Professional Development*, during spring 2016, the first time the

course was offered. Part of the online Masters in Teacher Leadership, the course enrolled 58 students, 44 (76%) of whom worked as teachers in K-12 schools. In the words of the syllabus, the course prepares teacher leaders to “identify, design, lead and evaluate professional development programs.” My responsibilities included facilitating discussions, assessing student work, responding to student questions and providing extra help as needed. Although I did not develop the course, I was asked to reflect on its design and offer suggestions for the future. Those reflections, along with a summary of our student evaluations, comprise this artifact. (Note: I structured my work with EDUC 774 as an independent study with Dr. Mouza, so the reflection essay also served as my final project for the course). This artifact demonstrates that I can effectively help practitioners understand the research behind professional development and can engage and support them in developing their own PD plans. (Note: an analysis of students’ PD plans is available in artifact j). It also shows my skills in self-evaluation.

Artifact h: Understanding by Design curriculum unit to teach graduate research assistants about PD evaluation. In EDUC 897, Curriculum Planning and Design, our final project in spring 2014 was to develop a curriculum unit. To make this assignment authentic, I developed a curriculum to orient the new graduate research assistants (GRAs) whom DERDC planned welcome in fall 2014. We anticipated a 100% turnover among GRAs and would quickly need to get our new students up to speed on understanding professional development and how to evaluate it. I structured the unit in the Understanding by Design (Wiggins & McTighe, 2006) format around the essential question: *What will it look like when PD is successful?* The artifact includes the unit itself (pp. 24-31) as well as a rationale and literature review

about professional development and PD evaluation, much of which I later integrated into my ELP proposal.

I used this curriculum but did not implement it fully as written. There were several reasons for this change. First, DERDC unexpectedly brought on one student for the summer. Since we had a compressed timeframe for orientation and a single student rather than a cohort, I could not use the full curriculum. (Luckily, this individual learned very quickly and was ready to go “into the field” even without all the preparatory experiences outlined in the curriculum). Later, we did welcome two GRAs in fall 2014 and experienced the opposite situation. The students had more basic developmental needs that dominated our time when they first arrived, and the curriculum was rather too ambitious. Still, I drew from it what I could (e.g., teaching and using the Guskey framework, reflecting on our own PD experiences, co-observing PD events) and both students did eventually get to the point of being able to independently collect, analyze, and report PD evaluation data.

Artifact i: Online PD toolkit. This artifact demonstrates teaching in a different, job-embedded context. It is found at:

<https://sites.google.com/a/udel.edu/online-pd-toolkit/?pli=1>

This was my final project for EDUC 818, Educational Technology Foundations. It is an online collection of resources and best practices to help people plan, design, implement, or evaluate online professional development (OPD). It could also be used by participants to figure out what to look for in high-quality OPD. I became interested in OPD when two client programs began offering online or opportunities. It was clear that while some principles of face-to-face PD translated, the online environment posed new challenges and opportunities for professional

developers and evaluators. I did not publicize this toolkit or teach it in any official way, but I drew on what I learned from it to design data collection instruments and more informally to advise project staff.

5. Increase understanding of what educators value in professional learning efforts and evaluations.

As described, I value practical knowledge and want my ELP to keep a foot in the real world of schools and teachers. This lever is about ensuring that university-based staff learn from practitioners and listen to their interests and their needs.

Artifact j: Exploration of evaluation use with former clients. My ELP rests on an assumption that studying professional development is one strategy for improving it, and that research/evaluation findings can and should be useful to practitioners. This artifact puts pressure on this assumption and investigates concepts of evaluation use. I will seek feedback from former evaluation clients, specifically staff from the three PD programs described under artifact e, above. I have stayed in touch with program managers from MADE CLEAR and DT3P and expect they will participate with this project. Because of their breadth, the Title II grants involved a more disparate group of staff members but I hope to gain their involvement as well. Drawing on literature about evaluation use and my knowledge about PD programs in general and these three in specific, I will develop interview protocols and/or other data collection instruments. I will encourage participants to be candid and specific about how they did or did not find my PD evaluations useful. The goals of this artifact are (1) to focus critical attention on my ELP's foundational logic (2) to model reciprocal relationships between evaluators and practitioners and (3) to push me to grow as a program evaluator.

Artifact k: Analysis of student PDPs from EDUC 774. This artifact examines the professional development plans (PDPs) that my students completed as final projects in EDUC 774. The overall purpose of the review is to examine whether and to what extent teacher PDPs demonstrate research-supported characteristics, and to provide insight into how a group of teacher leaders envision PD that meets their and their schools' needs. In November 2016, I informed EDUC 774 students about this research and invited them to participate. Using a passive process as approved by the IRB, all but one student gave consent. For alignment with the rest of my ELP, I decided to limit this analysis to plans created for Delaware elementary, middle and/or high schools (N=33). I developed a 10-item framework for analysis based on the Learning Forward Standards for Professional Learning and on other research syntheses about characteristics of effective PD (e.g., Desimone (2009)) and our course syllabus. I examined the plans for the presence and nature of each characteristic and identified interesting examples.

Table L2 Table of artifacts

Number	Artifact	Type	Audience	Description	Action Steps	Plans for IRB	Timeline	Status
A	Review of PD and PD eval frameworks	Literature review	PD developers, evaluators	See paragraph above	1. Clarify focus and subsections of review 2. Ask committee for suggested sources 3. Draft 4. Revise	None needed	Target for completion: July 31 2017	Not started
B	TLI study summary and recs	White paper	State policy makers and PD decision makers (e.g., district leaders)	See paragraph above	1. Write full TLI study 2. Excerpt relevant points for white paper 3. Draft 4. Obtain feedback from practitioners	TLI study will go through IRB (see	Target for completion: June 30 2017	Not started
C	Presentation @influence of org context on PD implementation	PPT	School leaders	See paragraph above	1. Review peer & comm feedback 2. Revise	None needed	Target for completion: March 31 2017	Pending feedback from committee
D	Study of the Schools That Lead	Research report	STL leadership,	See paragraph	1. Develop and pilot	Plan to submit	Target for completion:	In process

	Teacher Leadership Initiative	(qualitative)	fundors, Delaware DOE	above	interview protocols 2. Submit protocol to IRB 3. Attend STL meetings 1/31-2/2 4. Once IRB clearance recruit participants and collect data	protocol by 1/31/17	June 30, 2017	
E	Synthesis of three PD programs	Reflective essay	PD evaluators, designers	See paragraph above	1. Review final evaluation reports for three programs 2. Obtain committee support for developing review process 3. Draft & revise	Already done; evaluations in question have all been granted exempt status from the UD IRB already	Target for completion: April 30, 2017	Not started
F	Meta-evaluation of the SIIP grant	Evaluation report	DDOE	See paragraph above	1. Revise based on committee feedback	Already done; exempt status	Target for completion: February 28 2017	Complete pending committee feedback
G	Reflections from teaching EDUC 774	Reflection & course eval data	UD faculty esp. in MEd in TL	See paragraph above	None anticipated	None needed	Already complete	Complete
H	UBD curriculum unit about evaluating PD	Curriculum unit	Program evaluators, supervisors	See paragraph above	None anticipated	None needed	Already complete	Complete

I	Online PD toolkit	Website	PD developers, participants evaluators	See paragraph above	1. Review site for any needed updates 2. Revise	None needed	Target for completion: February 28 2017	Pending committee feedback
J	Exploration of evaluation use with former clients	Reflective essay	PD evaluators	See paragraph above	1. Contact project staff to discuss project, request participation 2. Seek input from committee about process, data collection 3. Research evaluation use to develop DCI	Depending on design I choose, this will probably need to go to the IRB with anticipated submission summer 2017	Target for completion: August 31 2017	Not started
K	Analysis of student PDPs	Content analysis	PD developers, evaluators, UD faculty	See paragraph above	1. Obtain committee feedback on draft	Already done; exempt status	Target for completion: February 28 2017	In process

References

- Albright, M. (2016, November 15). Delaware schools call for focus on disadvantaged children. *The News Journal*, A4.
- Blank, R.K. & de las Alas, N. (2010). Effects of teacher professional development on gains in student achievement: how meta analysis provides scientific evidence useful to education leaders. Society for Research on Educational Effectiveness.
- Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educational Researcher*, 33 (8), 3 – 15.
- Calvert, L. (2016). Moving from compliance to agency: What teachers need to make professional learning work. Oxford, OH: Learning Forward and NCTAF.
- Coldwell, M. & Simkins, T. (2011). Level models of continuing professional development evaluation: A grounded review and critique. *Professional Development in Education*, 37 (1), 143 – 157.
- Delaware Department of Education (2016a). Reimagining professional learning innovation grant application. Retrieved October 23, 2016 from: <http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/95/CIPD%20PL%20Grant%20FY16%20v4.pdf>
- Delaware Department of Education (2016a). State of Delaware school profile. Retrieved November 17, 2016 from: <http://profiles.doe.k12.de.us/SchoolProfiles/State/Account.aspx>
- Delaware Department of Education (January 3, 2014). Highly qualified teachers and improving teacher quality state grant program [Request for proposal].
- Desimone, L. (2011). A primer on effective professional development. *Phi Delta Kappan*, 92 (6), 668 – 671.
- Desimone, L. (2009). Improving impact studies of teachers' professional development: Towards better conceptualizations and measures. *Educational Researcher*, 38 (3), 181 – 199.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F. & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38 (4), 915 – 945.

- Bill & Melinda Gates Foundation (2014). Teachers know best: Teachers' views on professional development. Retrieved November 22, 2016 from: <http://k12education.gatesfoundation.org/wp-content/uploads/2015/04/Gates-PDMarketResearch-Dec5.pdf>
- Goldring, R., Taie, S., and Riddles, M. (2014). Teacher Attrition and Mobility: Results From the 2012–13 Teacher Follow-up Survey (NCES 2014-077). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved November 17, 2016 from <http://nces.ed.gov/pubsearch>.
- Guskey, T.R. (2002). Professional development and teacher change. *Teachers and Teaching*, 8 (3), 381-391.
- Guskey, T.R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership*, 59 (6), 45 – 51.
- Guskey, T. R. (2000). *Evaluating professional development*. Newbury Park, CA: Corwin Press.
- Hirsch, S. (2015, December 18). New bill offers good start on defining PD. *Learning Forward Blog*. Retrieved October 23, 2016 from: <https://learningforward.org/publications/blog/learning-forward-blog/2015/12/18/new-bill-offers-a-good-start-on-defining-pd>
<https://learningforward.org/publications/blog/learning-forward-blog/2015/12/18/new-bill-offers-a-good-start-on-defining-pd>
- Jacobs, H.H. (2012). *Curriculum 21: Essential education for a changing world*. Alexandria, VA: ASCD.
- Kennedy, A. (2005). Models of continuing professional development: a framework for analysis. *Journal of In-Service Education*, 31 (2), 235–250.
- Kennedy, A. (2014). Understanding continuing professional development: The need for theory to impact on policy and practice. *Professional Development in Education*, 40 (5) 688-697.
- Kennedy, M.M. (2016, in press). How does professional development improve teaching? *Review of Educational Research*. Advance online publication.

- King, F. (2014). Evaluating the impact of teacher professional development: An evidence-based framework. *Professional Development in Education*, 40 (1), 89 – 111.
- Layton, L. (2015, January 16). Majority of U.S. public school students are in poverty. *The Washington Post*. Retrieved November 17, 2016 from https://www.washingtonpost.com/local/education/majority-of-us-public-school-students-are-in-poverty/2015/01/15/df7171d0-9ce9-11e4-a7ee-526210d665b4_story.htmlhttps://www.washingtonpost.com/local/education/majority-of-us-public-school-students-are-in-poverty/2015/01/15/df7171d0-9ce9-11e4-a7ee-526210d665b4_story.html
- https://www.washingtonpost.com/local/education/majority-of-us-public-school-students-are-in-poverty/2015/01/15/df7171d0-9ce9-11e4-a7ee-526210d665b4_story.html
- Layton, L. (2015, August 4). Study: Billions of dollars in annual teacher training is largely a waste. *The Washington Post*. Retrieved November 17, 2016 from: https://www.washingtonpost.com/local/education/study-billionsof-dollars-in-annual-teacher-training-is-largely-awaste/2015/08/03/c4e1f322-39ff-11e5-9c2ded991d848c48_story.html
- Learning Forward (2015). Definition of professional development. Retrieved November 22, 2016 from: <https://learningforward.org/who-we-are/professional-learning-definition>
- Learning Forward (2011). Standards for professional development. Retrieved April 21, 2014 from <http://learningforward.org/standards>
- National Research Council (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- New Teacher Center (2013a). TELL Delaware results. Retrieved October 23, 2016 from: <https://telldelaware.org/results/report/80/1346#PD>
- New Teacher Center (2013b). How different educators perceive teaching conditions. Retrieved October 23, 2016 from: https://telldelaware.org/uploads/File/DE13_brief_comparisons.pdf
- Opfer, V.D. & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, 81 (3) 376 - 407.

- Penuel, W.R., Fishman, B.J., Yamaguchi, R. & Gallagher, L.P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, 44 (4), 921 – 958.
- Reeves, D. (2010). *Transforming Professional Development into Student Results*. Alexandria, VA: ASCD.
- Robertson-Kraft, C. & Hejlek, J. (2016). 2015 Delaware talent practices report. Philadelphia, PA: Operation Public Education. Retrieved November 17, 2016 from <http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/355/2015%20Delaware%20Talent%20Practices%20Report%20.pdf>
- [http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/355/2015Delaware Talent Practices Report .pdf](http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/355/2015DelawareTalentPracticesReport.pdf)
- Rodel Foundation (2015). 2015-16 Delaware public education at a glance. Retrieved November 17, 2016 from <http://www.rodelfoundationde.org/ataglance/#teacher-leader>
- Timperley, H. (2011). *Realizing the power of professional learning*. Maidenhead, England: Open University Press.
- TNTP (2015). The mirage: Confronting the hard truth about our quest for teacher development. Retrieved November 22, 2016 from: <http://tntp.org/publications/view/the-mirage-confronting-the-truth-about-our-quest-for-teacher-development>
- University of Delaware (2014). 2015 Documentation reporting form: Carnegie Community Engagement Classification (First Time Applicants). Retrieved January 6, 2017 from: <https://sites.udel.edu/engage/files/2016/08/University-of-Delaware-first-time-app-11elobr.pdf>
- University of Delaware (2015). Delaware will shine: A blueprint for a pre-eminent learner-centered research university. Retrieved January 6, 2017 from: <http://www1.udel.edu/research/delawarewillshine/DelawareWillShine-FINAL.pdf>

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007).
Reviewing the evidence on how teacher professional development affects
student achievement (Issues & Answers Report, REL 2007–No. 033).
Washington, DC: U.S. Department of Education, Institute of Education
Sciences, National Center for Education Evaluation and Regional Assistance,
Regional Educational Laboratory Southwest. Retrieved from
<http://ies.ed.gov/ncee/edlabs>

Appendix M

INSTITUTIONAL REVIEW BOARD LETTERS

The next pages include UD IRB exemption or approvals for the following activities conducted as part of this IRB:

- Title II Evaluation (Appendix A)
- Study of the Teacher Leadership Initiative (Appendix B)
- Evaluation of the SIIP grant (Appendix F)
- Analysis of Professional Development Plans (Appendix J)
- Analysis of evaluation use in PD programs (Appendix K)



DATE: June 13, 2014

TO: Joan Buttram
FROM: University of Delaware IRB

STUDY TITLE: [479610-2] Title II: Highly Qualified Teachers Evaluation

SUBMISSION TYPE: Amendment/Modification

ACTION: DETERMINATION OF EXEMPT STATUS
DECISION DATE: June 13, 2014

REVIEW CATEGORY: Exemption category # (1,2)

Thank you for your submission of Amendment/Modification materials for this research study. The University of Delaware IRB has determined this project is EXEMPT FROM IRB REVIEW according to federal regulations.

We will put a copy of this correspondence on file in our office. Please remember to notify us if you make any substantial changes to the project.

If you have any questions, please contact Nicole Farnese-McFarlane at (302) 831-1119 or nicolefm@udel.edu. Please include your study title and reference number in all correspondence with this office.



RESEARCH OFFICE

210 Hulihan Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: February 28, 2017

TO: Joan Buttram
FROM: University of Delaware IRB

STUDY TITLE: [1017951-1] Study of the Teacher Leadership Initiative

SUBMISSION TYPE: New Project

ACTION: APPROVED
APPROVAL DATE: February 28, 2017
EXPIRATION DATE: February 27, 2018
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # (6,7)

Thank you for your submission of New Project materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.



DATE: April 29, 2014

TO: Joan Buttram
FROM: University of Delaware IRB

STUDY TITLE: [601820-1] Evaluation of the SIIP grant

SUBMISSION TYPE: New Project

ACTION: DETERMINATION OF EXEMPT STATUS
DECISION DATE: April 29, 2014

REVIEW CATEGORY: Exemption category # (3)

Thank you for your submission of New Project materials for this research study. The University of Delaware IRB has determined this project is EXEMPT FROM IRB REVIEW according to federal regulations.

We will put a copy of this correspondence on file in our office. Please remember to notify us if you make any substantial changes to the project.

If you have any questions, please contact Nicole Farnese-McFarlane at (302) 831-1119 or nicolefm@udel.edu. Please include your study title and reference number in all correspondence with this office.

cc:



DATE: October 17, 2016

TO: Hilary Mead
FROM: University of Delaware IRB (HUMANS)

STUDY TITLE: [970074-1] Analysis of Professional Development Plans

SUBMISSION TYPE: New Project

ACTION: DETERMINATION OF EXEMPT STATUS
DECISION DATE: October 17, 2016

REVIEW CATEGORY: Exemption category # (1)

Thank you for your submission of New Project materials for this research study. The University of Delaware IRB (HUMANS) has determined this project is EXEMPT FROM IRB REVIEW according to federal regulations.

We will put a copy of this correspondence on file in our office. Please remember to notify us if you make any substantial changes to the project.

If you have any questions, please contact Nicole Farnese-McFarlane at (302) 831-1119 or nicolefm@udel.edu. Please include your study title and reference number in all correspondence with this office.

cc:



RESEARCH OFFICE

210 Hulihan Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: March 30, 2017

TO: Hilary Mead
FROM: University of Delaware IRB (HUMANS)

STUDY TITLE: [1050450-1] Analysis of evaluation use in PD programs

SUBMISSION TYPE: New Project

ACTION: APPROVED
APPROVAL DATE: March 30, 2017
EXPIRATION DATE: March 29, 2018
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # (6,7)

Thank you for your submission of New Project materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.