International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 12 ǁ December 2022 ǁ PP. 194-208

Structuring Secondary Mathematics Teacher Preparation Through a Professional Development Framework

Christopher R. Rakes[1],Jon Saderholm[2],Sarah B. Bush[3],Margaret J. Mohr-Schroeder[4],Robert N. Ronau[5],Michele Stites[1]
*[*1]Department of Education, University of Maryland Baltimore County, Baltimore, MD, USA*
[2]Department of Education Studies, Berea College, Berea, KY, USA
[3]College of Community Innovation and Education, University of Central Florida, Orlando, FL, USA
[4]College of Education, University of Kentucky, Lexington, KY, USA
[5]College of Education and Human Development, University of Louisville, Louisville, KY, USA
Corresponding Author:Christopher Rakes, rakes@umbc.edu

Abstract
This study examined the use of the PrimeD framework to improve secondary (Grades 7-12) mathematics teacher preparation. The study used design-based research through a four-year treatment-only mixed methods triangulation design. Data sources were program documents and assessments, focus groups, interviews, and field experience observations. The Reformed Teaching Observation Protocol with Performance Descriptors (RTOP+) measured candidates' teaching quality at three time points during the student teaching internship. As the program incorporated more features of PrimeD, supervisors and mentor teachers increasingly took on leadership roles and each subsequent cohort demonstrated stronger growth on the RTOP+.

Keywords:*Mathematics Education, Secondary Education, Teacher Preparation, Design-Based Research, Improvement Science*

Date of Submission: 02-12-2022 Date of acceptance: 13-12-2022

I. INTRODUCTION

The present study examined the use of the Professional Development: Research, Implementation, and Evaluation framework (PrimeD; [1-2]) to improve mathematics teacher preparation. Strengthening the knowledge and skills of Science, Technology, Engineering, and Mathematics (STEM) teachers is nationally recognized as a high priority [3-5]. Teacher preparation programs are considered one of the greatest leverage points for long-term improvement in teacher performance and retention of effective teachers [6-8]. Teacher preparation programs in the U.S. struggle to meet their maximum potential because their components are often fragmented, leading to artificial divisions between various types of knowledge, theory, and practice [9]. An important way to address such fragmentation is to provide a common frame of reference for quality teaching [10].

Compounding such systemic issues in teacher preparation programs, teacher candidates (referred to as ―candidates" throughout) themselves hold many misconceptions about teaching that must be addressed to fully prepare them for the profession. For example, candidates often consider teaching to be primarily an interpersonal, affective activity rather than as a profession requiring particular skills and knowledge [11]. Such misconceptions are especially problematic for teacher preparation because incomplete understanding of content and pedagogy and lack of connections between content and pedagogy can inhibit quality teaching [6]. Furthermore, candidates sometimes view teacher preparation as making demands that are unrealistic and beyond the scope of normal teacher expectations [12]. Candidates often believe their preparation programs promote broad directives with little instruction on how to carry out those directives. For example, candidates frequently assert that courses tell them to adjust lessons to accommodate special needs but seldom show them how to do so [13].

The PrimeD framework, the innovation of the present study, was designed to provide a coherent structure to professional development (PD) activities and may also address the type of fragmentation often found in teacher preparation programs (e.g., divisions between subject matter and pedagogy, as in [9]). PrimeD organizes PD into four phases: Design, Implementation, Evaluation, and Research. Based on early field trials, PrimeD has demonstrated flexibility to also provide structure for STEM education PD programs, helping them to address both systemic and individual challenges [14]. PrimeD Phase I Design can readily be tailored to directly address the types of candidate misconceptions described by Bangel et al. [13], Fajet et al. [11], and

Pollock et al. [12]. Using networked improvement communities (NICs), PrimeD positions participants as researchers in their own classrooms to investigate well-defined problems of practice and to report those findings and refine their classroom innovations in PD sessions. The inclusion of evaluation and research in the program design is a unique feature of PrimeD that directly connects measures and assessments to the program design and implementation.

1.1 Research Questions

The present study explored the implementation of PrimeD in a single teacher preparation program. The study examined a single secondary (Grades 7-12) mathematics teacher preparation program through a four-year longitudinal mixed methods design. The questions that guided the study were:

1. In what ways did PrimeD change the secondary mathematics program?
2. In what ways did PrimeD reposition stakeholders (e.g., classroom teachers, field experience supervisors, teacher candidates, faculty) as leaders, partners, and co-learners in the program?
3. To what degree did PrimeD improve teacher candidates' preparation to implement research-based teaching strategies?

Teacher preparation is often viewed as initial professional development (e.g., [12, 15]), and many aspects of effective PD have been well-studied (e.g., [16-17]). While the literature provides much grounding for a variety of effective activities within a mathematics teacher preparation program (e.g., [18-19]), the existence of many gaps in the research knowledge base limits the ability of the field to make conclusions about systemic characteristics that make teacher preparation effective [5, 20-22]. PrimeD is a comprehensive framework that supports systemic as well as more fine-grained innovations in a program.

II. Conceptual Framework: PrimeD

The PrimeD framework was initially developed through a systematic review of literature [23] and through the evaluation of a state-wide PD program in the U.S. [2]. PrimeD applies the principles of improvement science to PD [24-26] and synthesizes the recommendations of several PD frameworks (e.g., [16-17]). PrimeD organizes PD into four phases that work in a cyclic nature and occur iteratively throughout a PD program (Figure 1).

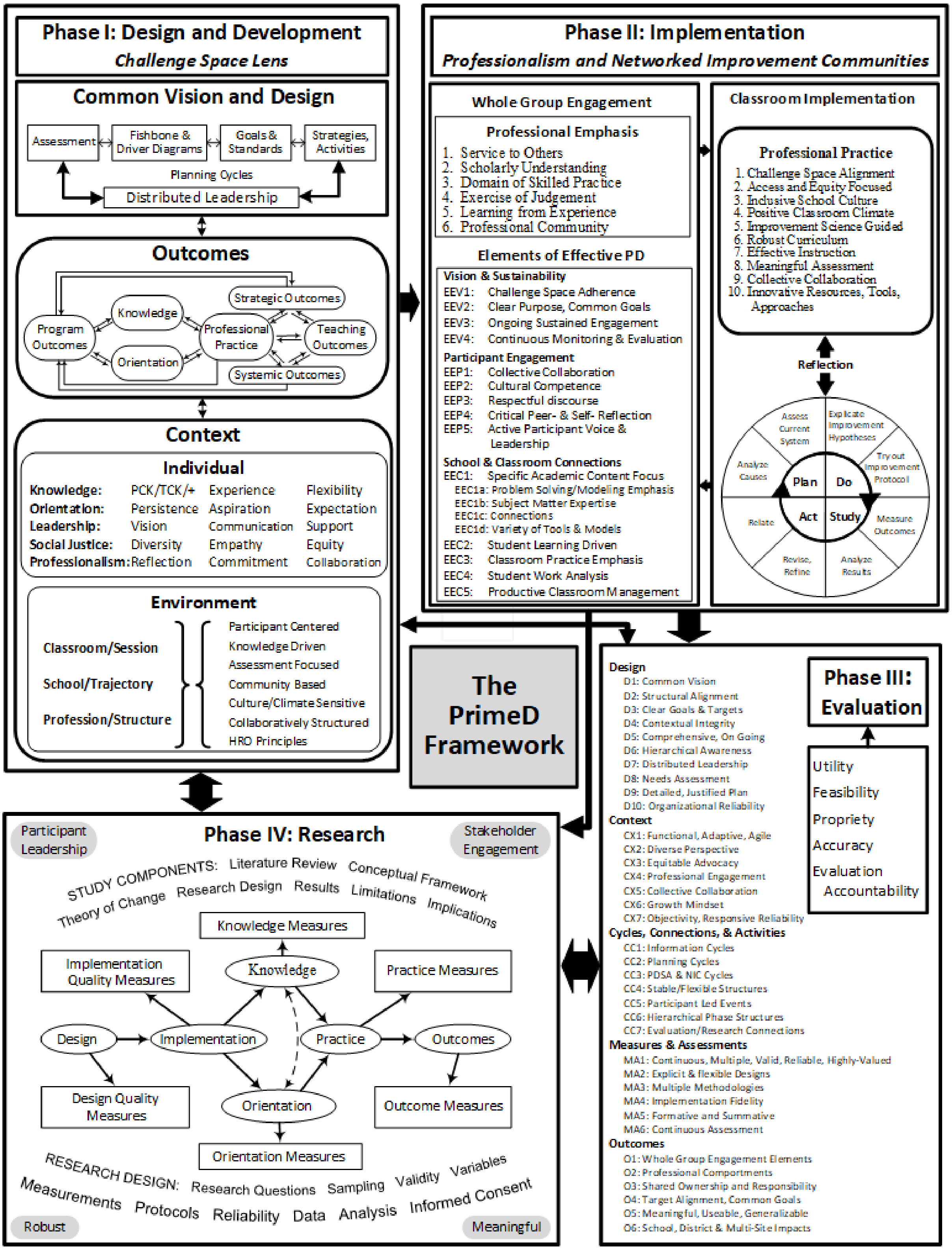

Figure 1. Model of the PrimeD framework.

2.1 Phase I Design and Development

Phase I focuses on the construction of a challenge space by all stakeholders. In teacher preparation, stakeholders include groups such as university faculty, classroom mentor teachers, field experience supervisors, and candidates. Stakeholders come together to map out the challenge space—an explicit description of the needs, vision, goals, targets, and strategies for meeting the challenges being addressed by and faced within the program [24]. The challenge space is more than a list of obstacles or difficulties; it embodies the program‘s call to action to improve entry-level professional practice and expresses a pragmatic vision of the potential for systemic and systematic change [1]. Phase I is foundational to every other phase in PrimeD. It goes far beyond simply planning PD (coursework and field experiences in a teacher preparation program). Each course, session, and field experience should be purposefully aligned to the challenge space. But perhaps more importantly in teacher preparation, structural supports are needed to bind course and field experiences together into a coherent system, promoting collaboration and focusing on classroom practice [24, 28-30]. The development of the challenge space begins the process of engaging all partners productively, as recommended by the Association of Mathematics Teacher Educators [20]. The strategies planned in Phase I are informed by the elements of effective PD and professional practice characteristics defined in Phase II.

2.2 Phase II Implementation

Phase II Implementation is divided into two structural components, *whole group engagement* and *classroom implementation*. Networked Improvement Communities (NICs) and Plan-Do-Study-Act (PDSA) cycles provide the overarching structure to connect whole group activities to classroom practice [30]. NICs are intentionally designed social organizations that are: 1) focused on a common aim, 2) guided by a deep understanding of the problem and a shared approach to solve it, 3) disciplined by the methods of improvement research to develop, test, and refine interventions; and 4) organized to accelerate interventions and effectively integrate them into the field. In a NIC, stakeholders are encouraged to work together to develop a common vision. To be true to the NIC construct, participants must be engaged as active partners rather than recipients in all activities. The problems in a teacher preparation NIC focus on connecting research and theory to classroom practice, hereafter ―problems of practice." A problem of practice must be observable and actionable, focused on instruction, and connected to the overall improvement strategy. For example, teachers may be concerned that students rely too much on memorization, make computation errors, and lack understanding why mathematical procedures work or when to apply them [31].

2.2.1 Whole Group Engagement

Within whole group engagement, professionalism undergirds the problem of practice, and the PrimeD elements of effective PD provide structure for group activities.Professionalism demands that teachers embrace lifelong learning of their craft, never settling for the status quo [32]. Professionalism is characterized by service to others, scholarly understanding, skilled practice, judgment under uncertainty, learning from experience, and responsible professional communities [33]. The teaching profession focuses on student learning, facility with research-based practices, and collective collaboration [16, 20, 34-35]. The PrimeD elements of effective PD were synthesized from several studies about PD (e.g., [16-17, 36]). These elements align closely to the vision of effective PD described by the National Council of Teachers of Mathematics (NCTM) [37-38], for example being sustained over time, focusing on important mathematics, and providing meaningful opportunities for teachers to collaborate.

The NIC structure creates opportunities for teacher candidates to engage in professional collaboration with each other, their classroom mentor teachers, faculty, and other school leaders (e.g., instructional coaches, administrators) as they explore connections between research, theory, and practice. PrimeD in teacher preparation recognizes that professionalism requires attention to the growth of all participants, including classroom mentor teachers. The NIC structure provides opportunities for classroom mentor teachers to engage in the program as leaders and resources for knowledge and experience, aligning to recommendations by McAleer [39] for facilitating mentor professional growth.

2.2.2 Classroom Practice

Professional practice demands explicit attention to equity [38]. In whole group engagement, discussions about equity might include systemic issues such as student and teacher tracking. In classroom practice, the focus is on developing robust mathematics lessons that open the conceptual space for all students (e.g., increased student communication, multiple representations, climate of respect;[40]) and providing instructional supports to ensure the success of all students (e.g., additional time [38]). Equitable, meaningful assessment is an indispensable component of classroom practice, for example aligning expectations with assessments, providing multiple forms of assessment, giving high-quality and consistent feedback, consciously attending to the influence of biases and assumptions about student ability, and distinction of assessment from grades [41-43].

A Plan-Do-Study Act (PDSA) cycle is a tightly planned cycle of classroom practice in which a purpose/goal is clearly defined, an activity is proposed and implemented, evidence is collected to determine the result, and the activity is refined based on that evidence. In some cases, the evidence may indicate that the goal should be adjusted [24-25, 44]. The Plan stage begins during whole group engagement, in which participants in the NIC establish a common problem of practice, aligned to the challenge space. Planning consists of defining the problem and change, making predictions about what will happen as a result of the change, and forming a plan for assessing the change [25]. During the Do stage, participants carry out the change and collect data. During the Study stage, participants analyze the data, compare the results to their predictions, and document insights to be used in the next cycle. The Act stage is about deciding what to do next based on what was learned. For example, the change idea might be abandoned, adjusted, or expanded [25]. Through multiple cycles, initial ideas are refined into effective practices and eventually systemic change. Using PDSA cycles in teacher preparation situates mentors and teacher candidates as researchers in their classrooms as recommended by Philippou et al. [45]. PDSA cycles can be thought of as iterative action research focused on a common problem of practice and an agreed-upon change. Such connections between research and practice are especially important for transferring research into practice [46-48]. Feedback from practice to academic research situates the practitioner as an integral component of the development of knowledge in the field, recognizing the possibility that practitioner knowledge based on experience may run deeper than academic theory [49]. PDSA cycles in student teaching may provide a more targeted approach for helping candidates encounter multiple ways in which an effective teaching practice can be enacted (as recommended by Gainsburg [27]), allowing them to create a more general model of effective teaching at an accelerated pace.

2.3 Phase III Evaluation

Because PrimeD focuses on engaging teachers as professionals, evaluation is an integral component. As professionals, teachers participate in establishing what is best practice [50]. Engaging candidates in evaluation lays a foundation for professionalism throughout their careers. Evaluation cycles in PrimeD include feedback mechanisms to the challenge space (Phase I). The five PrimeD evaluation categories (Design, Context, Cycles & Activities, Measures & Assessments, and Outcomes) act as lenses to structure both formative and summative evaluation. PrimeD evaluation was developed to align with the five Joint Committee on Standards for Educational Evaluation (JCSEE) program evaluation standard recommendation categories: utility, feasibility, propriety, accuracy, and evaluation accountability [51]. PrimeD evaluation goes beyond determining whether a particular innovation worked, providing continuous cycles of feedback to provide a roadmap to improvement. The rapid, formative feedback cycles from challenge space to implementation to evaluation to challenge space (Phases I to II to III to I) align with recommendations by Bryk et al. [25] and Rogers [52]. Such feedback cycles necessitate having instruments and measures that are well-aligned to the challenge space [26].

2.4 Phase IV Research

Evaluation and research are inter-related yet distinct [53]. Evaluation and research both involve investigation and inquiry, and similar methods and data sources are used. Evaluation, however, is generally context-specific whereas research seeks generalizability [53]. Evaluation questions address process (fidelity of implementation) and whether the outcomes are reached while research questions examine the factors that may have influenced the degree to which outcomes are reached.

The relationship between research and evaluation may be context dependent [54-55]. That is, in some projects, research and evaluation may be completely separate, in others they may intersect, and in still others one may be a subset of the other. Moreover, the relationship between evaluation and research may evolve during a project or may change as directed from the research/evaluation process itself. The research component of the framework has a hierarchical nature: research on an overall project, on various components within a project (e.g., the process and effects of the various NICs), on characteristics of teacher preparation, and on effective teaching practice.

If the intention is to prepare teachers to enter a professional community, then treating them as professionals must be integral to their preparation program so that monitoring quality and developing knowledge for teaching becomes the norm [50]. Engaging in such reflective, collaborative work during teacher candidacy lays a strong foundation for future professional growth. Teachers regularly carry out research in their classrooms [56] and seek out research that is directly applicable to the classroom [57]. Teachers may at the same time think of ―research” as a hands-off activity with little connection to the classroom [57]. Methods such as design-based research are especially useful to support partnerships between researchers and practitioners with a goal of generating outcomes that are both practical and contribute to theory [56]. Viewing research as a seamless component of PD adds access, richness and complexity to the process and has been shown to improve professional learning outcomes for teachers (e.g., [58-60]).

Teachers conduct research as a normal function of their practice; that is, they test and evaluate their approach to teaching every day, seeking causal explanations for outcomes they observe. But these types of efforts are often contextually limited. PrimeD offers a structure that can make these normal situated research activities robust and useful for a larger audience by connecting implementation (Phase II) and research (Phase IV) activities. While implementing PD innovations in Phase II‘s PDSA cycles, teachers create research questions from their classrooms and provide critical insight into questions about the PD itself. PDSA cycles are intended to be rapid trials, the results of which are initially shared in a NIC. Results are generalized in Phase IV when they are shared with the larger group to be tried and vetted to determine what works and does not work for desired outcomes under various conditions and why [61]. In Phase IV, the program may use a variety of approaches and designs to structure the generalization of results beyond specific classroom contexts. Connecting teacher research into the overall PD research agenda can lead to richer insights and avoid having their findings lost to the larger community [62]. PrimeD supports the development of strong, meaningful partnerships between teachers and researchers, enabling further study about how research impacts classrooms directly and thus increases the ability for such efforts to enhance the knowledge of the field.

The inclusion of Phase IV in teacher preparation indicates an intentionality to the goal of laying the groundwork for candidates to engage in professional research as teachers. Through the NIC and PDSA cycles, candidates observe mentor activities, ask questions, engage collaboratively, and develop the necessary foundations for contributing to the knowledge base. Mentors, supervisors, and faculty help guide the conclusions candidates make about their classroom research trials, helping to hone their professional judgment.

In PrimeD Phase IV, plans, activities, and results are in a constant state of study in an open structure shared with all participants and stakeholders. For example, the challenge space includes planning and information cycles and is subject to modification if such change is warranted by the results. On the other hand, not every participant and stakeholder may wish (or are able) to participate in every cycle, but they have a choice to participate in those cycles that are meaningful and/or important to them.

III. Methods

The present study used design-based research [63] to follow a secondary (Grades 7-12) mathematics teacher preparation program through a treatment-only, longitudinal, mixed methods triangulation design [64-65]. The study hypothesized that PrimeD can provide a common frame of reference for quality teaching, strengthen linkages to practice, and enhance the degree to which field experiences inform teacher preparation programs.

3.1 Context

The study took place within a single university teacher preparation program located in an urban area in the mid-Atlantic region of the U.S. The university is a high research, doctoral degree granting public institution with an enrollment of approximately 14,000 students. The secondary (Grades 7-12) mathematics program includes undergraduate, graduate, and fifth-year program pathways to becoming a teacher. The program typically has approximately 40 students at various stages of the program with anywhere from four to eleven students completing the program each year.

The secondary education program follows a cohort model in which all teacher candidates complete two phases of student teaching ―internship." Candidates are typically placed with the same classroom mentor teacher for both phases. Phase I is in the fall semester while Phase II is in the spring. Phase I of the internship requires the candidates to be in their placement for two days a week for a minimum of 20 full days. Phase II of the internship is the full-time student teaching experience and requires a minimum of 80 days. The mathematics methods course is taught in the fall in conjunction with Phase I. An internship seminar course for the secondary program meets three times in the fall and 11 times in the spring (all secondary content areas). Of the 11 seminar meetings in the spring (during Phase II), four sessions are content-specific, meeting approximately once per month.

3.2 Procedures

The study gradually incorporated components of PrimeD each year. In Year 1, the program challenge space was developed and refined through multiple discussions with candidates, field experience supervisors, classroom mentor teachers, and university faculty. In Years 2-4, the program incorporated and refined the PrimeD elements of effective PD, NICs, PDSA cycles, and evaluation and research cycles.

3.3 Measures

Data sources consisted of program documents and assessments, focus groups, interviews, and field experience observations. In Year 2, the Reformed Teaching Observation Protocol (RTOP;[40]) was added to the internship to better align the feedback and measures to the challenge space. RTOP is widely used as quantitative and qualitative tool in STEM education (e.g., [66-67]). With 25 indicators rated on a 0 to 4-point scale (0 = not descriptive of the lesson, 4 = very descriptive of the lesson), RTOP was developed to assess the degree to which teaching is reformed; that is, with a focus on student empowerment, critical thinking, active learning, and a problem-solving approach [68]. Reformed teaching goes beyond classroom practice to include a classroom culture that supports change and moves away from lecture toward constructivism (Sawada et al., 2002). RTOP was designed to focus exclusively on characteristics of reformed teaching rather than broader concepts of good teaching (e.g., lesson closure; [69]). The indicators were based on mathematics and science education research and national recommendations and standards, including NCTM's *Principles and Standards of School Mathematics*[70] and National Science Education Standards [71-72]. Expert panels were used to assess content validity [69]. Construct validity was measured using exploratory factor analysis: the subscales were found to account for approximately 71% of the variance [40, 69]. Inter-rater reliability was measured between raters on overall scores with an average r-squared value of .954 for mathematics and physics classes [40, 69]. RTOP scores also showed strong predictive validity with student achievement in mathematics using correlation between average RTOP scores and normalized achievement gain scores, $r = .94$ [40, 69].

The RTOP was enhanced with descriptors to provide stronger distinctions between performance levels for each indicator to enhance validity of the scores and to enhance inter-rater reliability [73]. Intraclass correlation coefficients (ICC) were interpreted using Cicchetti's criteria [74]: Excellent = 0.75 – 1, Good = 0.60 – 0.74. Two university supervisors rated teacher candidate lessons with the RTOP + Performance Descriptors (RTOP+). Supervisors double coded at least one lesson independently in each year the RTOP+ was used (Years 2-4). ICC was measured for reliability across the 25 RTOP+ items and was either good or excellent each year: Year 2, ICC = .738; Year 3, ICC =.740; Year 4, ICC = .828.

3.4 Analytic Methods

This mixed methods triangulation design consisted of concurrent qualitative and quantitative analyses. Program documents were collected and analyzed each year to document the types of changes implemented in the program. Focus groups and interviews were administered each year to gather participant perspectives and conduct member checks. Three lessons were scored with RTOP+ each year: Lesson 1 in late November/early December (end of Phase I internship), Lesson 2 in late January/early February (beginning of Phase II internship, full time student teaching), and Lesson 3 in late April/early May (end of Phase II internship). Unstructured interviews were conducted at the end of Phase II internship. Thick description was used to identify emergent themes. Repeated measures ANOVA was used to analyze RTOP+ scores.

IV. Results

Broadly stated, the present study examined the extent to which the use of PrimeD as a framework for teacher preparation strengthened candidates' ability to operate as professionals in their early career. As design-based research, the implementation of PrimeD was iterative, with each year building on lessons learned from the previous year.

We begin by presenting how PrimeD was implemented in the teacher preparation program over the four years of the study and how the implementation changed the program and the positionality of stakeholders within the program (Research Questions 1 and 2). Following the description of PrimeD implementation over time, the results continue by describing how equity, a major focus of the program's challenge space, was addressed across the four years. We conclude with an analysis of candidates' RTOP+ scores (Research Question 3).

4.1 Program Development from PrimeD Implementation Over Time

The incorporation of PrimeD into the teacher preparation program influenced the overall goals of the program, connections between coursework, field experiences, and assessments. Ongoing evaluation led to major changes (e.g., incorporating NICs into the program) at the beginning of each academic year with minor changes and tweaks (e.g., format of whole group meetings) occurring throughout the year. The advisory board provided feedback biweekly.

4.1.1 Changes in Year 1: Initial PrimeD Implementation

In Year 1, a program challenge space (e.g., needs, goals, targets, contexts, strategies; PrimeD Phase I) was developed and refined among stakeholders (e.g., faculty, supervisors, mentor teachers, candidates). Program goals focused on teacher knowledge, teacher orientation, teacher practice, and student outcomes. Although goals were listed for each category, Year 1 efforts emphasized teacher practice goals for the classroom learning culture, especially use of culturally relevant pedagogy, active engagement of students, inquiry-based instruction, focus on conceptual understanding and quantitative reasoning, connections to authentic contexts, and questioning and assessment strategies.

PDSA cycles were added to the program, but the templates from teacher PD research proved to be too complex and unwieldy. Candidates struggled to complete the forms, and program structures were not sufficiently in place to support the process.

Stakeholders discussed the challenge space throughout the year and developed fishbone diagrams to map out potential contributing factors to a problem. Three overarching issues were investigated: the degree to which teacher candidates apply mathematics content knowledge to their secondary classroom teaching, the degree to which teacher candidates felt the internship was overwhelming, and the range of pedagogies demonstrated during student teaching (theory→practice; Figure 2).

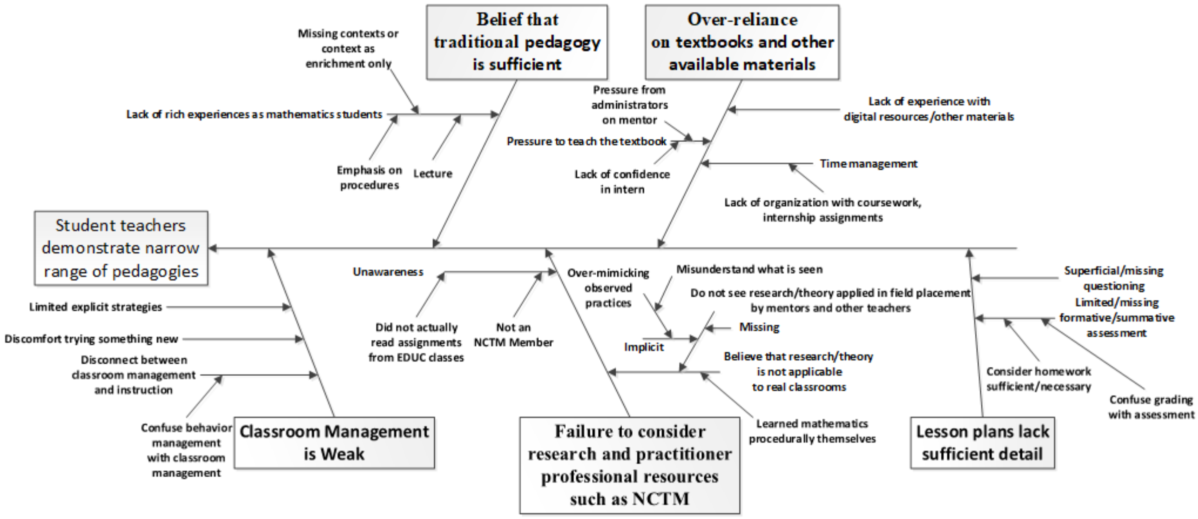

Figure 2. Example of fishbone diagram developed in Year 1.

Candidates helped develop the fishbone diagrams and discussed possible solutions. For example, to address the overwhelming nature of the internship experience, class time was set aside to discuss personal care strategies, work samples with rubric scores, and ways to adjust due dates to meet program goals without piling multiple assignments on at the same time. By the end of Year 1, driver diagrams had been developed. Figure 3 provides an example diagram for how the program was intended to support the development of reformed teaching and equitable teaching practices.

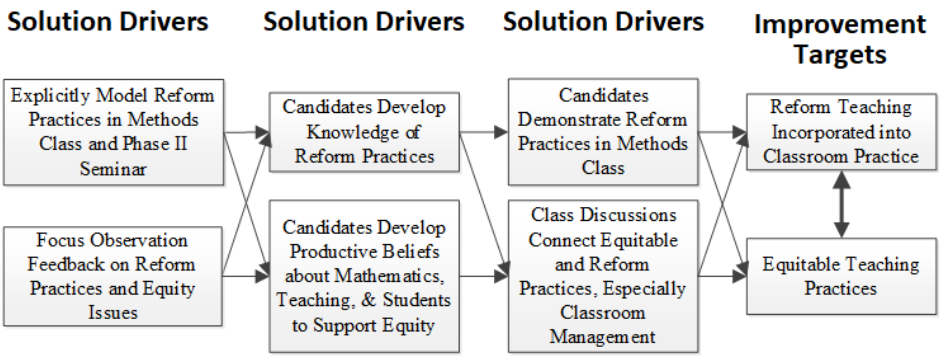

Figure 3. Driver diagram of program improvement targets developed during Year 1.

4.1.2 Evolution of PrimeD Implementation in Years 2 and 3

In Year 2, the challenge space was further refined by faculty, program alumni, mentors, and supervisors based on Year 1 data and feedback, and an evaluation plan was developed. As part of the evaluation plan, field experience assessments were revised to include the RTOP+. The rubric and its performance descriptors were also integrated into coursework as an instructional tool to focus candidates' pedagogical development on inquiry-based strategies that promoted divergent mathematical thinking (e.g., Rows 4, 17), mathematics learning through scientific inquiry (e.g., Rows 12, 15), collaborative learning and student communication (e.g., Rows 16, 20), and student leadership (e.g., Rows 13, 19). For example, Row 12 promotes the use of scientific inquiry to learn mathematics and states, ―Students made predictions, estimations, and/or hypotheses and devised means for testing them (collecting and analyzing data) Conjecture evident." A lesson at Level 4 would demonstrate, ―The students explicitly stated what they expected the outcomes of the activity were going to be and devised a means to test the prediction and collect data." At Level 3, ―The students stated in general or vague terms what they expected the outcomes of the activity were going to be and devised a means to test the prediction and collect data. Some teacher guidance." At Level 2, ―Students made predictions, but these predictions were followed up by classroom discussion and teacher directed explanations. Methodology provided." And at Level 1, ―Students were given a hypothesis to test or discuss. Step by step process. No prediction. (Cookbook activity)." The full rubric is provided as an online supplementary material.

The program‘s driver diagram was revised to incorporate evaluation feedback and to tighten connections between reformed teaching and equity (Figure 4). The new driver diagram guided the program in both Years 2 and 3.

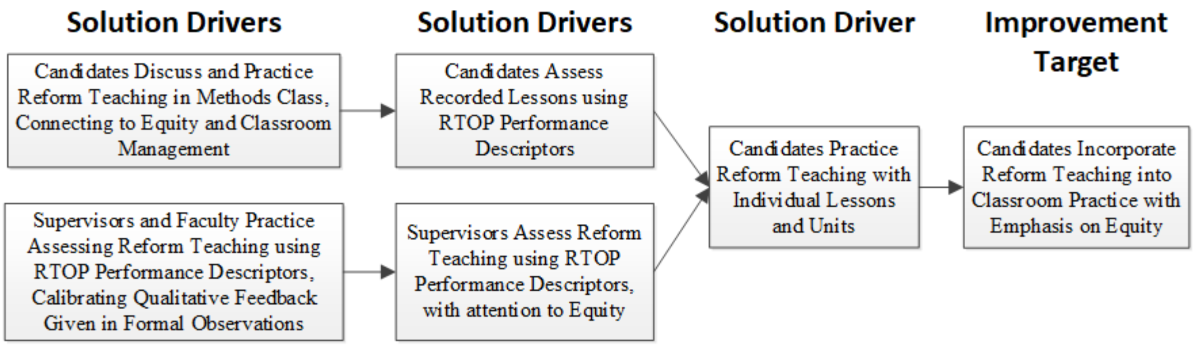

Figure 4. Driver diagram of improvement target for Years 2 and 3.

The methods course was revised to better address the PrimeD elements of effective PD, especially participant engagement: active participant voice and leadership. Reading assignments were categorized by the component of the challenge space being addressed (e.g., equity, specific content pedagogy, assessment). Articles were chosen for each category based on the strength of connections to research, theory, and practice with an emphasis on professional practitioner journals such as NCTM‘s *Mathematics Teacher* and *Teaching Mathematics in the Middle School*. Teams of three or four candidates were assigned a reading category rather than a specific article. The teams chose an article from within that category, presented it to the class, and then presented a demonstration of how to apply that reading to their professional practice. A focus group at the end of Phase I internship reported that candidates felt a greater sense of autonomy and empowerment in the new methods class structure and that the program was generally better focused on their needs. They said that the reading presentations made learning more enjoyable and fostered creativity, community, and leadership. Moreover, they felt prepared for Phase II internship and were excited to try out strategies they had learned in class from each other. They recognized where they were and where they were going in the teacher preparation process. During a focus group, one candidate stated, ―I mean before when I looked at it, I was like ‗What?‘ But now I think when I‘m doing it in Phase II, I will know specifically what I‘m trying to ask you if I have a question about it. I‘m so glad this was not introduced to me only in Phase II." They reported that they had already tried out some strategies during Phase I but expected to have improved agency to try out more in Phase II as they began taking over their mentors‘ classes.

PDSA cycles were aligned to the timetable of the field experience observations conducted by the university supervisors, approximately once per month with three cycles in Phase I internship and four cycles in Phase II. The PDSA template was revised into a series of questions or prompts for each phase of the cycle. Candidates completed the questions before and after each observation, maintaining the document as a running record of their professional growth.

An internship debrief was instituted at the end of Phase II. These debrief sessions were focused on discussing feedback from the internship portfolio and observations, with an emphasis on preparing for the first year of teaching (e.g., strengths, areas for improvement, professional goals). In Year 2, these were conducted as one-on-one sessions between candidates and faculty. In Year 3, supervisors also participated.

4.1.3 Addition of NICs to Program in Year 4

In Year 4, the program was restructured around NICs and PDSA cycles, building on all innovations from Years 1-3. Classroom mentor teachers, supervisors, candidates, and faculty participated in the NIC meetings. Candidates and mentors guided the discussions, resulting in a focus problem and innovation: using spiral instruction [75] to address gaps between foundational knowledge and new content. Spiral instruction became the central focus of their internship year and provided a lens through which they viewed pedagogical innovations. Candidates explicitly attempted spiral instruction in their classrooms in collaboration with their mentor teachers, collected data, and reported their results at subsequent whole group NIC meetings. In the first few meetings, candidates expanded and refined their understanding of spiral instruction. As their experience grew, NIC discussions grew to include implementing spiral instruction for various lessons and topics, developing valid student outcome measures, improving assessment techniques, and conducting research on their practice. These NIC and PDSA activities were incorporated into the preparation program‘s existing course and field experience structures. The full implementation of NICs and PDSA cycles affected every other aspect of PrimeD implementation (e.g., methods course discussions and topics, field experience observations, participant interaction with program evaluation measures, elements of effective PD: vision and sustainability, participant engagement, school and classroom connections). The NIC structure provided a forum for discussing and debating ideas, and PDSA cycles focused candidate efforts to integrate ideas from the NIC meetings into their classes. At the end of Phase I internship during a methods class, candidates assessed a recorded lesson with the RTOP performance descriptors. They stated during a focus group that assessing the video was the first time they felt like they understood what the RTOP performance descriptors meant. They wished that it had been done earlier in the semester and that it had been done as part of a NIC meeting so that they could discuss with mentors.

NIC whole group discussions about spiral instruction led to pedagogical calibration across all candidates. A student learning objective (SLO) project was integrated into the NIC: the cycles of whole group discussions and classroom implementation were a natural fit for the SLO project, so the group decided to do a single overall project with three sub-groups working together. These sub-groups were formed based on common content, grade levels, and school communities. This collaboration increased the generalizability of their classroom-level results, providing insights into the use of spiral instruction in multiple grade levels, schools, districts, and content domains. During the Phase II internship, school districts and university shifted to online learning due to COVID-19. Data collection on their project was curtailed. But in the process of running their project, they had established communication structures that were used to support one another in the move to online. This pivot also meant that the final NIC meeting was held online. In their post-internship debriefs, candidates consistently referred to spiral instruction as the foundation for their professional growth for the following year. A new logic model emerged from the evaluation feedback cycle to guide future years of the program (Figure 5).

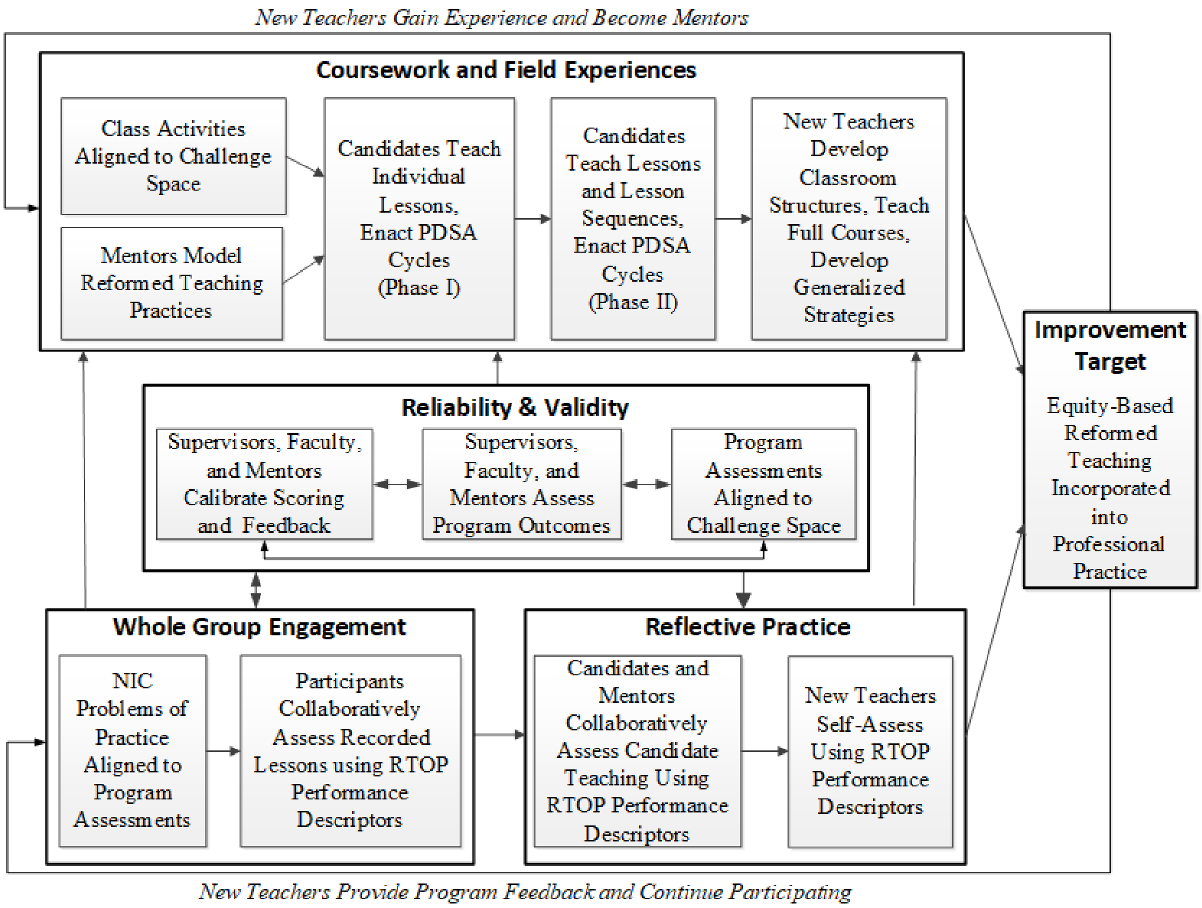

Figure 5. Final logic model for program.

This new logic model recognizes that the improvement target goal is not realized when candidates complete the program but rather when, as new teachers, they develop generalized strategies to incorporate reformed teaching into their professional practice. Such a target emphasizes the need to have a coherent program from candidacy through the early teaching career.

4.2 The Role of Equity in the Program Challenge Space

The program challenge space was founded on the notion that teaching is only good to the degree that it is equitable. Equity was a multilevel concern for every aspect of the preparation program (―equity for candidates"). In Year 1, equitable teaching was treated as a separate improvement target, but in Years 2-4, it was treated as an integral component of reformed teaching (―equity for students").

4.2.1 Equity at the Program Level

Shrewsbury [76] identified empowerment, community, and leadership as foundational tenets of equitable teaching. Candidates were empowered to take on a leadership role in the program, altering the topics addressed in class sessions, taking ownership for research and theory connections to practice (e.g., studying, presenting, demonstrating, debating). In focus group discussions, candidates agreed that the methods class structures such as the reading presentations empowered them. One candidate stated, ―The class was student centered to me, and it gave us a lot of autonomy." Strong interpersonal bonds were forged as participants shared ideas, aspirations, experiences, and concerns. Explicit attention was given to building a learning community, in which participants took responsibility for each other as well as their own learning. For example, candidates took on a leadership role in the program, altering the topics addressed in class sessions, taking ownership for research and theory connections to practice (e.g., studying, presenting, demonstrating, debating). Candidates practiced assessing sample lesson plan sequences, video-recorded lessons, and feedback on K-12 student work using the program assessment rubrics (e.g., RTOP+; EdTPA rubrics). Candidates stated that they found these sessions to be extremely helpful with their own struggles in addressing program expectations and requested an expansion of these practice sessions at all subsequent whole-group meetings. This feedback was treated as a program evaluation and was incorporated into the program challenge space through a revision of program strategies. Candidates were involved in the revision process to increase transparency and empower them to continue providing feedback.

4.2.2 Equity in Candidate Practice

Equity and inclusion were treated as their own topics and as foundational components of every other topic throughout the program. NCTM's *Principles to Actions* and *Catalyzing Change in High School Mathematics*[37-38] along with AMTE's *Standards for Preparing Teachers of Mathematics*[20] served as guiding documents along with research and theory about culturally relevant pedagogy, social justice in mathematics, teacher-student relationships, the role of classroom structures in promoting equity, teacher expectations, ethnomathematics, groups historically marginalized in mathematics classrooms, and general and content-specific pedagogy for supporting struggling students. The relationship of spiral instruction, conceptual understanding, and equity were explored throughout each year. Strategies for promoting positive mathematics identities and valuing students' diverse perspectives were emphasized.

4.3 Reformed Teaching Scores

Scores for the RTOP+ were collected three times for each cohort, once during Phase I internship (semester before full time student teaching) and then at the beginning and end of Phase II (full time student teaching). Table 1 provides the descriptive statistics for each cohort.

Table 1RTOP+ Descriptive Statistics

Cohort	Time	N	Min	Mean	SD	Max
2017-18	1	8	58	61.9	3.3	66
	2	8	54	58.1	4.7	67
	3	**8**	**60**	**62.7**	**2.4**	**66**
2018-19	1	4	42	58.0	18.5	84
	2	4	60	70.0	7.8	78
	3	**4**	**46**	**72.8**	**18.9**	**87**
2019-20	1	9	28	50.1	10.7	62
	2	8	44	57.8	9.1	70
	3	**8**	**74**	**77.6**	**3.0**	**82**

Note. Time 1 = end of Phase I (mid-November/early December). Time 2 = beginning of Phase II (late January/early February). Time 3 = End of Phase II (late April/early May). Time 3 is bold to emphasize the highest cohort means.

A repeated measures ANOVA (RM-ANOVA) was conducted on the three scores by cohort. Mauchly's Test of Sphericity indicated that the data did not violate the sphericity assumption, $W(\text{df} = 2) = .948$, $p > .5$. The RM-ANOVA for the Time by Cohort interaction was statistically significant, $F(4,30) = 8.038$, $p < .001$.

Because significance was found in the RM-ANOVA, a post hoc mean difference analysis was conducted using a Bonferroni adjustment for multiple comparisons. This analysis found statistical differences between Times 1 and 3 ($p < .001$) and between Times 2 and 3 ($p = .003$), but not between Times 1 and 2 ($p = .116$). A line plot (Figure 6) shows that the growth from Time 1 to Time 3 increased with each successive cohort, 0.8 points in 2017-18, 14.8 points in 2018-19, and 27.5 points in 2019-20.

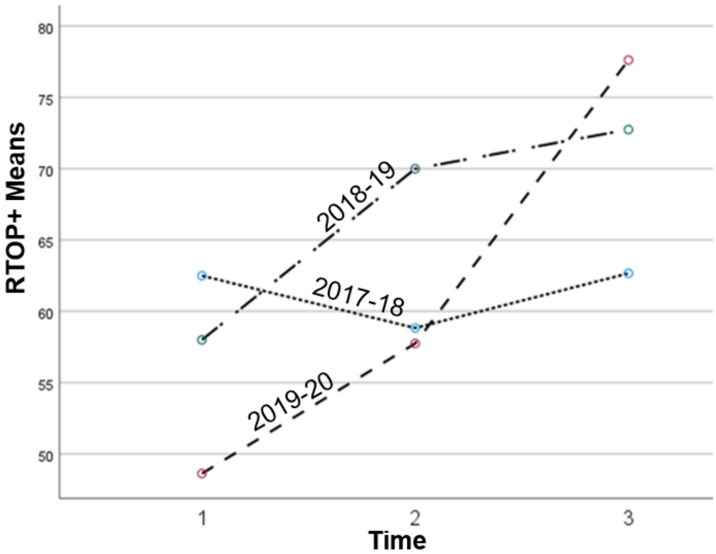

Figure 6. RTOP+ means for each cohort by time.

The slight decline in 2017-18 from Time 1 to 2 warranted further discussion with candidates, especially since they indicated in their Phase I focus group that they expected to be able to enact more reformed teaching during Phase II. In a follow up focus group discussion, they stated that they had felt largely overwhelmed at the beginning of Phase II as they started to assume the leadership of their mentors' classes. They said that they had largely gone into ―survival mode" and fell back on more familiar traditional teaching strategies. They suggested two ways to better support candidates in the transition to Phase II: (a) have more explicit discussions about how to manage the transition to full time teaching to avoid being overwhelmed, (b) spend more time at the end of the methods course discussing foundational, general strategies for teaching in a more reformed way, and (c) refocus candidates on reformed teaching strategies prior to Time 2 lesson. Their advice was applied in the 2018-19 cohort, which showed steadier gains over time. In 2019-20, their advice was also applied within the context of the NIC meetings: focusing on reformed teaching from the outset of Phase II became the problem of practice at the end of Phase I. With mentors more heavily involved in the discussions throughout the year, candidates reported feeling more supported and empowered to teach in a reformed way.

V. Conclusions

The use of PrimeD led to substantial changes to the structure of the secondary mathematics teacher preparation program. The program goals, outcomes, and strategies (challenge space) were better articulated with all stakeholders. Feedback cycles from stakeholders led to refinements in the challenge space, especially strategies. The addition of the RTOP+ as a formative assessment led to stronger evaluation cycles that guided stakeholder understanding of the challenge space and whole group discussions, an example of PrimeD Phase 3 informing Phase 1, which in turn informed subsequent enactments of Phase 2. All three cohorts outperformed national norms for experienced teachers (48 to 50 points for secondary mathematics, [40]) by Time 3, demonstrating a strong ability to teach in a reformed way within their student teaching contexts given sufficient levels of support.

The methods course revisions provided candidates with multiple opportunities to take on a leadership role in the program. The reading presentations helped strengthen connections between theory and practice. In conjunction with the NIC problems of practice, candidates were better able to understand and focus on explicit teaching methods in their professional growth.

The incorporation of NICs shifted the role of mentors from solely field experience partners to professional learning partners with faculty, supervisors, and candidates. These partnerships guided the discussion topics during whole group activities and bolstered the degree to which candidates could engage in PDSA cycles. Supervisors, mentors, and faculty were better able to collaborate on providing consistent, focused feedback to candidates related to the NIC problems of practice and reformed teaching in general.

VI. Limitations

The present study explored the promise of using PrimeD framework as a guiding structure for secondary mathematics teacher preparation programs. With only a single program as a sample, the study is limited in its generalizability to programs with dissimilar contexts. The COVID-19 pandemic led to shutdowns of face-to-face instruction at both the university and partner schools, limiting the comparability of subsequent data. The study did not follow candidates into their early career, so the degree to which they continued incorporating reformed teaching into their practice is unclear.

VII. Discussion

The results of this study demonstrated that a professional development framework such as PrimeD can enhance a secondary mathematics teacher preparation program. The use of PrimeD led to stronger relationships between faculty, supervisors, and mentors as partners and professionals. Candidates better understood the types of instructional practices being targeted by the goals of the program and demonstrated a strong foundation with which to begin their careers. Follow up studies may examine the generalizability of PrimeD in teacher preparation by engaging with multiple institutions and contexts. They may also include an early career component to better ascertain how such a program could support early career teachers and leverage their experiences to better support subsequent teacher candidate cohorts. An early career program may be vital to helping new teachers develop generalized strategies for reformed teaching, as recommended by Gainsburg [27].

Teaching is a profession, and effective preparation for teachers requires more than motivation and knowledge. The process of preparing teachers to be professionals requires well-focused initiatives that establish mutually agreed-upon goals and help bridge the divide between theory and practice [17, 25]. Professionals share leadership and collaborate to improve the nature of their craft [50]. By engaging candidates and mentors in such collaborations through NICs, the program helps candidates begin their teaching careers with a strong sense of what it means to be a professional. PDSA cycles provide a structure for enacting theory in the classroom. The use of assessments as an instructional tool empowers candidates to guide their own learning.

Enacting PrimeD in teacher preparation was a gradual yet drastic shift from traditional preparation programs. Participants (e.g., faculty, supervisors, candidates, mentors) learned daily how the way the program applied the framework could and needed to evolve to meet new or newly understood challenges. We encourage the use of PrimeD as outlined here to guide the work of teacher preparation. Using PrimeD as an explicit model to support research-based practices can lead to meaningful structural improvement of teacher preparation and increase the degree to which candidates are well-prepared to enter the field.

REFERENCES

[1]. Rakes, C. R., Bush, S. B., Mohr-Schroeder, M. J., Ronau, R. N., & Saderholm, J. (2017). Making teacher PD effective using the PrimeD framework. *New England Mathematics Journal, 50*, 52-63.

[2]. Saderholm, J., Ronau, R. N., Rakes, C. R., Bush, S. B., & Mohr-Schroeder, M. (2017). The critical role of a well-articulated conceptual framework to guide professional development: An evaluation of a state-wide two-week program for mathematics and science teachers. *Professional Development in Education, 43*, 789-818. https://doi.org/10.1080/19415257.2016.1251485

[3]. National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (2010). *Rising above the gathering storm revisited: Rapidly approaching Category 5.* Washington, DC: National Academies Press.

[4]. National Research Council. (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future.* Washington, DC: National Academies Press.

[5]. National Research Council. (2010). *Preparing teachers: Building evidence for sound policy.* Washington, DC: National Academies Press.

[6]. Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives, 8* (1), 1-44. http://epaa.asu.edu/ojs/article/view/392/515

[7]. Ginsberg, R., & Rhodes, L. K. (2003). University faculty in partner schools. *Journal of Teacher Education, 54*, 150-162. https://doi.org/10.1177/0022487102250308

[8]. U.S. Department of Education. (2018). Teacher professional and career development [website]. https://www.ed.gov/oii-news/teacher-professional-and-career-development

[9]. Ball, D. L. (2000). Bridging practices: Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education, 51*, 241-247. https://doi.org/10.1177/0022487100051003013

[10]. McKay, S., & Silva, E. (2015). *Improving observer training: The trends and challenges* [Issue Brief]. Stanford, CA: Carnegie Foundation for the Advancement of Teaching. https://www.carnegiefoundation.org/resources/publications/improving-observer-training-the-trends-and-challenges/

[11]. Fajet, W., Bello, M., Leftwich, S. A., Mesler, J. L., & Shaver, A. N. (2005). Pre-service teachers‘ perceptions in beginning education classes. *Teaching and Teacher Education, 21*, 717-727. https://doi.org/10.1016/j.tate.2005.05.002

[12]. Pollock, M., Bocala, C., Deckman, S. L., & Dickstein-Staub, S. (2016). Caricature and hyperbole in preservice teacher professional development for diversity. *Urban Education, 51*, 629-658. https://doi.org/10.1177/0042085915581714

[13]. Bangel, N. J., Enersen, D., Capobianco, B., & Moon, S. M. (2006). Professional development of preservice teachers: Teaching in the super Saturday program. *Journal for the Education of the Gifted, 29*, 339-361. https://doi.org/10.1177%2F016235320602900305

[14]. Bush, S. B., Mohr-Schroeder, M. J., Cook, K. L., Rakes, C. R., Ronau, R. N., & Saderholm, J. (2020). Structuring integrated STEM education professional development: Challenges revealed and insights gained from a cross-case synthesis. *Electronic Journal for Research in Science and Mathematics Education, 24*(1), 26-55. https://ejse.southwestern.edu/article/view/19940

[15]. Darling-Hammond, L. (2005). Teaching as a profession: Lessons in teacher preparation and professional development. *Phi Delta Kappan, 87*, 237-240. https://doi.org/10.1177%2F003172170508700318

[16]. Desimone, L. M. (2009). Improving impact studies of teachers' professional development: toward better conceptualizations and measures. *Educational Researcher, 38*, 181-199. https://doi.org/10.3102%2F0013189X08331140

[17]. Loucks-Horsley, S., Stiles, E., Mundry, S., Love, N., & Hewson, P. (2010). *Designing professional development for teachers of science and mathematics. (*3rd ed.). Thousand Oaks, CA: Corwin.

[18]. Feldman, Z., & Roscoe, M. B. (2018). Encouraging teachers to make use of multiplicative structure. *Mathematics Teacher Educator, 7*, 60-85. https://doi.org/10.5951/mathteaceduc.7.1.0060

[19]. I, J.-Y., & Stanford, J. (2018). Preservice teachers' mathematical visual implementation for emergent bilinguals. *Mathematics Teacher Educator, 7*, 8-33. https://doi.org/10.5951/mathteaceduc.7.1.0008

[20]. Association of Mathematics Teacher Educators. (2017). *Standards for preparing teachers of mathematics*. http://amte.net/standards

[21]. Taylor, P. M., & Ronau, R. N. (2006). Syllabus study: A structured look at mathematics methods courses.*AMTE Connections, 16* (1), 12-15. https://amte.net/connections

[22]. Wilson, S. M., Floden, R. E., Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations. A research report prepared for the U.S. Department of Education* (Document R-01-3).Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.

[23]. Driskell, S. O., Bush, S. B., Ronau, R. N., Niess, M. L., **Rakes, C. R.**, & Pugalee, D. (2016). Mathematics education technology professional development: Changes over several decades. In M. L. Niess, S. O. Driskell, & K. F. Hollebrands (Eds.), *Handbook of research on transforming mathematics teacher education in the digital age* (pp. 107-136). Hershey, PA: IGI Global. http://hdl.handle.net/11603/14415

[24]. Bryk A. S., Gomez L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. T. Hallinan (Ed.), *Frontiers in Sociology in Education* (pp. 127-162). Dordrecht, Netherlands: Springer Netherlands.

[25]. Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. (2015). *Learning to improve. How America's schools can get better at getting better.* Cambridge, MA: Harvard Educational Publishing.

[26]. Yeager, D., Bryk, A., Muhich, J., Hausman, H., & Morales, L. (2013). *Practical measurement*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

[27]. Gainsburg, J. (2012). Why new mathematics teachers do or don't use practices emphasized in their credential program. *Journal of Mathematics Teacher Education, 15*, 359-379. https://doi.org/10.1007/s10857-012-9208-1

[28]. Levine, A. (2006). *Educating school teachers.* Washington, DC: The Education Schools Project. http://edschools.org/pdf/Educating_Teachers_Report.pdf

[29]. Luft, J. A., Diamond, J. M., Zhang, C., & White, D. Y. (2020). Research on K-12 STEM professional development programs: An examination of program design and teacher knowledge and practice. In C. C. Johnson, M. J. Mohr-Schroeder, T. J. Moore, & L. D. English (Eds.), *Handbook of Research on STEM Education* (Ch. 29, 14 pages). New York, NY: Routledge. https://doi.org/10.4324/9780429021381

[30]. Martin, W. G., & Gobstein, H. (2015). Generating a networked improvement community to improve secondary mathematics teacher preparation: Network leadership, organization, and operation. *Journal of Teacher Education, 66*, 482-493. https://doi.org/10.1177%2F0022487115602312

[31]. National Council of Teachers of Mathematics. (2021). *Innov8: Identifying your problem of practice.* Reston, VA: Author. https://www.nctm.org/innov8pop/

[32]. Collins, J. (2001). *Good to great: Why some companies make the leap…and others don't.* New York, NY: Harper Collins.

[33]. Shulman, L. S. (1998). Theory, practice, and the education of professionals. *The Elementary School Journal, 98*, 511-526. https://doi.org/10.1086/461912

[34]. Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York, NY: Simon and Schuster.

[35]. Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

[36]. Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*, 4-15.

[37]. National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all.* Reston, VA: Author.

[38]. National Council of Teachers of Mathematics. (2018). *Catalyzing change in high school mathematics: Initiating critical conversations.* Reston, VA: Author.

[39]. McAleer, S. D. (2008). Professional growth through mentoring: A study of experienced mathematics teachers participating in a content-based online mentoring and induction program. *Dissertation Abstracts International-A, 69*(08). (UMI No. 3319930)

[40]. Sawada, D., Piburn, M. D., Judson, E., Turley, J., & Falconer, K. (2002). Measuring reform practices in science and mathematics classrooms: The Reformed Teaching Observation Protocol. *School Science and Mathematics, 102*, 245-253. https://doi.org/10.1111/j.1949-8594.2002.tb17883.x

[41]. Copur-Gencturk, Y., Cimpian, J. R., Lubienski, S. T., & Thacker, I. (2020). Teachers' bias against the mathematical ability of female, Black, and Hispanic students. *Educational Researcher, 49*, 30-43. https://doi.org/10.3102%2F0013189X19890577

[42]. Porter, A. C., Smithson, J., Blank, R., & Zeidner, T. (2007). Alignment as a teacher variable. *Applied Measurement in Education, 20*, 27-51. https://doi.org/10.1080/08957340709336729

[43]. Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education* (ED414305). Washington, DC: Council of Chief State School Officers. ERIC. https://files.eric.ed.gov/fulltext/ED414305.pdf

[44]. Langley, G. J., Moen, R. D., Nolan, K. M., Nolan, T. W., Normal, C. L., & Provost, L. P. (2009). *The improvement guide: A practical approach to enhancing organizational performance* (2nd ed.). San Francisco, CA: Jossey-Bass.

[45]. Philippou, S., Papademetri-Kachrimani, C., & Louca, L. (2015). ‗The exchange of ideas was mutual, I have to say': Negotiating researcher and teacher ‗roles' in an early years educators' professional development programme on inquiry-based mathematics and science learning. *Professional Development in Education, 41*, 382-400.

[46]. Jones, K., & O'Brien, J. (2014). Introduction: Professional development in teacher education: European perspectives. In K. Jones & J. O'Brien (Eds.), *European perspectives on professional development in teacher education* (pp. 1-6). New York, NY: Routledge.

[47]. Sabah, S. A., Fayez, M., Alshamrani, S. M., & Mansour, N. (2014). Continuing professional development (CPD) provision for science and mathematics teachers in Saudi Arabia: Perceptions and experiences of CPD providers. *Journal of Baltic Science Education, 13*, 91-104. http://oaji.net/articles/2015/987-1437062396.pdf

[48]. Timperley, H. S. (2011). *Realizing the power of professional learning*. Maidenhead, ENG: Open University Press.

[49]. Hawkins, D. (1966). Learning the unteachable. In L. S. Shulman & E. R. Keisler (Eds.), *Learning by discovery: A critical appraisal* (pp. 3-12). Chicago, IL: Rand McNally & Company.

[50]. Shulman, L. S. (2004). *The wisdom of practice: Essays on teaching, learning, and learning to teach* (1st ed.). San Francisco, CA: Jossey-Bass.

[51]. Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

[52]. Rogers, P. J. (2021, April 15). Why do we need more real-time evaluation? Better Evaluation. https://www.betterevaluation.org/en/blog/why-do-we-need-more-real-time-evaluation

[53]. Chyung, S. Y. (2015). Foundational concepts for conducting program evaluations. *Performance Improvement Quarterly, 27*, 77-96. https://doi.org/10.1002/piq.21181

[54]. Rogers, P. (2014, May 9). Week 19: Ways of framing the difference between research and evaluation. Better Evaluation: Sharing information to improve evaluation.http://www.betterevaluation.org/en/blog/framing_the_difference_between_research_and_evaluation

[55]. Rogers, P. J., & Peersman, G. (2014). Developing a research agenda for impact evaluation in development. *IDS Bulletin*, 45(6), 85-99. https://doi.org/10.1111/1759-5436.12115

[56]. Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher, 41*, 16-25. https://doi.org/10.3102%2F0013189X11428813

[57]. Drill, K., Miller, S., & Behrstock, E. (2012). Teachers' Perspectives on Educational Research (EJ1057457). *Brock Education, 23*, 3-17. ERIC. https://files.eric.ed.gov/fulltext/EJ1057457.pdf

[58]. Atay, D. (2008). Teacher research for professional development. *ELT Journal, 62*, 139-147.https://doi.org/10.1093/elt/ccl053

[59]. Elliott, J. (1990). Teachers as researchers: Implications for supervision and for teacher education (ED293831). *Teaching and Teacher Education,6*, 1-26. ERIC. https://files.eric.ed.gov/fulltext/ED293831.pdf

[60]. Savoie- Zajc, L., & Descamps- Bednarz, N. (2007). Action research and collaborative research: Their specific contributions to professional development. *Educational Action Research,15*, 577-596. https://doi.org/10.1080/09650790701664013

[61]. Kirkwood, M., & Christie, D. (2006). The role of teacher research in continuing professional development. British Journal of Educational Studies, 54, 429-448. https://doi.org/10.1111/j.1467-8527.2006.00355.x

[62]. Cochran-Smith, M., & Lytle, S. (1990). Research on teaching and teacher research: The issues that divide. *Educational Researcher,19*, 2-11. https://doi.org/10.2307/1176596

[63]. The Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 5-8*.https://doi.org/10.3102/0013189X032001005

[64]. Creswell, J.W. and Plano Clark, V.L. (2011).*Designing and conducting mixed methods research*. (2nd Ed.).Thousand Oaks, CA: Sage.

[65]. Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences.*Thousand Oaks, CA: Sage.

[66]. Amolins, M. W., Ezrailson, C. M., Pearce, D. A., Elliott, A. J., & Vitiello, P. F. (2015). Evaluating the effectiveness of a laboratory-based professional development program for science educators. *Advances in Physiology Education, 39*, 341–351. https://doi.org/10.1152/advan.00088.2015

[67]. Enderle, P., Dentzau, M., Roseler, K., Southerland, S., Granger, E., Hughes, R., & Saka, Y. (2014). Examining the influence of RETs on science teacher beliefs and practice. *Science Education, 98*, 1077–1108. https://doi.org/10.1002/sce.21127

[68]. Anderson, R. D., Anderson, B. L., Varanka-Martin, M. A., Romagnano, L., Bielenberg, J., Flory, M., Mieras, B., & Whitworth, J. (1994). *Issues of curriculum reform in science, mathematics and higher order thinking across the discipline* (ED368064). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement. ERIC. https://ia802607.us.archive.org/13/items/ERIC_ED368064/ERIC_ED368064.pdf

[69]. Piburn, M., & Sawada, D. (2000). *Reformed Teaching Observation Protocol (RTOP): Reference manual.* (ACEPT Technical Report No. IN00-3). Phoenix, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers. http://www.public.asu.edu/~anton1/AssessArticles/Assessments/Biology%20Assessments/RTOP%20Reference%20Manual.pdf

[70]. National Council of Teachers of Mathematics. (2000). *Principles and standards of school mathematics.* Reston, VA: Author.

[71]. National Research Council. (1996). *National Science Education Standards.* Washington, DC: National Academy Press.

[72]. National Research Council. (2000). *Inquiry and the National Science Education Standards.* Washington, DC: National Academy Press.

[73]. Williams, T., Singer, J., Krikorian, J., Rakes, C. R., & Ross, J. (2019). Measuring pedagogy and the integration of engineering design in STEM classrooms. *Journal of Science Education and Technology, 28*, 179-194. https://doi.org/10.1007/s10956-018-9756-y

[74]. Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluation normed and standardized assessment instruments in psychology. *Psychology Assessment, 6*, 284–290. https://doi.apa.org/doi/10.1037/1040-3590.6.4.284

[75]. Gibbs, B. C. (2014). Reconfiguring Bruner: Compressing the spiral curriculum. *Phi Delta Kappan, 95*, 41-44. https://doi.org/10.1177/003172171409500710

[76]. Shrewsbury, C. M. (1993). What is feminist pedagogy? *Women's Studies Quarterly, 3-4*, 8-15. https://www.jstor.org/stable/40003432