Program:	Indicate which legislative appropriation this plan addresses. ***Math Achievement*** *This evaluation will focus on three programs implemented through this appropriation:* ● ***Focus on Algebra*** *focuses on helping educators, administrators, and instructional coaches who develop pedagogy and deliver instructional practices to students in grades 6-9 to emphasize algebraic thinking and concepts to prepare students for Algebra 1.* ● ***NUMeROS*** *builds math content knowledge for elementary teachers to provide them with the resources needed to build strong conceptual knowledge and mathematical reasoning in elementary students.* ● ***HQIM Implementation*** *supports educators, administrators and instructional coaches with grants to support implementation of high-quality instructional materials implementation, necessary to support use of core materials for on-grade level standards aligned instruction.*
Problem Statement:	Use this space to write a succinct problem statement the program is attempting to address. *Statewide math achievement is holding at 24% for the past two academic years. Increasing the use, coherence, and consistency of strong instructional practices in math courses is the most direct way to impact student outcomes in math across the state.*
Overarching Goals:	Use this space to list three to five overarching goals for the program. 1. *Improve educators' math instruction practices and knowledge* 2. *Increase the number of students who pass math courses* 3. *Increase student academic achievement in math* To strengthen monitoring and accountability, PED will define clear thresholds for success related to each key outcome. Thresholds and effect sizes will be determined once baseline data is reviewed and will be revisited regularly to ensure they remain appropriate.
Underlying Assumptions and Research Base: ***Why do you think this program will work?***	Use this space to include a brief summary regarding whether the program has been shown to work and whether the program is classified as evidence-based, research-based, promising, or does not yet have rigorous research. For definitions of these terms, see below and Section 6-3A-3 NMSA 1978. *Both the Focus on Algebra and NUMeROS programs have been evaluated internally, highlighting the potential benefits of each program. The Charles A. Dana Center out of the University of Texas at Austin evaluated their implementation of the Focus on Algebra program during the 2023-24 school year and found promising results that participating teachers and learners experienced gains in knowledge and self-efficacy regarding implementing*

	improved math instructional methods in their classrooms and schools. Additionally, the evaluation found that teachers were implementing the skills they learned in their classrooms. A review of a micro-credential course series utilizing NUMeROS found that nearly two-thirds of participants strongly agreed that the course deepened their knowledge about instructional approaches that increase student conceptual understanding of mathematics. *HQIM Implementation has a broad research basis, However, LEAs have not received support with implementing HQIM alongside curriculum-based professional learning. This program draws on a research-based evidence base but requires evaluation to understand the impact of HQIM implementation alongside professional learning in the New Mexico context.* *While these programs do not yet have a rigorous research base, this evaluation will provide an opportunity to build a research base that could benefit both New Mexico and the field at large.*

Logic Model

Building a logic model helps clearly articulate what a program aims to accomplish. It shows the relationships between program resources, activities, outputs, and expected outcomes, and serves as a framework for planning, implementation, and evaluation. In the logic model below, fill in the boxes to identify who is responsible for implementation, the activities each actor will carry out, how you will measure whether activities are implemented as intended, and the expected short- and long-term outcomes. See the Institute for Education Sciences for more information about logic models.

Actors	Activities	Outputs	Outcomes
List the entities will be responsible for implementing the program (one row per actor). *Examples: PED, school districts and charter schools, schools, educators, evaluators.*	List the specific activities each actor will be responsible for carrying out during implementation. *Examples: Awarding funding, providing technical support, collecting and cleaning data, training staff.*	List the measures used to determine whether activities are being implemented as intended. *Examples: Number of participants, attrition rate, number of times an activity was carried out, engagement metrics.*	List the short-term and long-term measures that may be used to provide evidence that the program is making a difference. **Short-Term**: Benefits for participants. **Long-Term**: Benefits for entire program. *Examples: Knowledge, skills, proficiency rates, graduation rates, chronic absence rates, additional outcomes for participants.*
LEAs	**Services and Supports for Students** ● Support curricular and pedagogical changes resulting from participation in programs **Program Administration and Monitoring** ● Execute grant agreements, data sharing agreements, and other programmatic documents (i.e. Memorandum of Agreements) with PED ● Identify students taught by educators participating in programs ● Identify school liaison to provide information to PED ● Engage in regular check-ins with PED and vendors ● Collect and review data to ensure student eligibility and progress, including meeting programmatic requirements, and share with PED ● In partnership with PED and vendors, support educators' ongoing engagement with program professional learning and support	**Services and Supports for Students** ● # of students taught by educators participating in these programs **Program Administration and Monitoring** ● PED Assurances/Memorandum of Agreements executed ● Data from students taught by participating educators shared with PED ● Student outcome data shared with PED	**Short-Term Outcomes** ● Teachers improve math instructional practices and knowledge ● Increase in number of students who pass math courses
Educators	**Program Administration and Monitoring** ● Attend summer trainings for programs ● Participate in ongoing professional learning and support events during the school year ● Participate in information gathering (conducted by vendors) regarding change in classroom practices	**Program Administration** ● # of professional learning and support events attended	

	● Implement data measurement tools and report results to LEA, PED and vendors **Services and Supports for Students** ● Implement content from trainings in curriculum development and instructional methods		
PED	**Program Administration and Monitoring** ● Select LEAs participating in each of the programs ● Award funding for HQIM program ● Execute Assurances/Memorandum of Agreements with participating LEAs ● Provide oversight of consultants delivering professional learning and monitoring classroom-level program implementation ● In partnership with LEAs, support educators' ongoing engagement with program professional learning and support	**Program Administration and Monitoring** ● # of LEAs and schools participating in each program ● Amount of funding provided for participants in HQIM program	**Long-Term Outcomes** ● Increase in student academic achievement in math.
Vendors	**Program Administration and Monitoring** ● Plan and deliver summer trainings ● Provide ongoing professional learning throughout the year ● Monitor quality of professional learning experiences ● Monitor classroom-level program implementation ● Evaluate implementation of program elements ● Gather information from educators regarding changes in classroom practices ● Gather information from students of participating educators regarding engagement with math	**Program Administration and Monitoring** ● # of educators participating in summer trainings ● Survey feedback from summer training participants ● Survey, interview, and focus group data from teachers and administrators about change in classroom practices ● Survey data from students taught by program participants about engagement with math ● Walkthrough and observation information on classroom implementation of programs	

Program Evaluation Plan and Description of Methods

A program evaluation plan provides a roadmap for determining whether a program is achieving its expected outcomes. It should include information about the target population, data collection methods, and planned data analysis. If needed, refer to the World Bank pre-analysis plan checklist for additional guidance.

Target Population and Sampling	● What is the target population of the program? (For example, is the program run at a classroom level, a school level, or a districtwide level? Is the program predominantly for a specific demographic or at-risk group?) o Both NUMeROS and Focus on Algebra are run at a school level, with school teams consisting of one school leader or instructional coach to deliver professional learning, one instructional leader able to make school decisions, and at least 80 percent of school staff responsible for mathematics instruction. The HQIM program runs at the district level, focusing on supporting district-level staff in adopting strong HQIM selection processes. o The programs will be targeting school level personnel responsible for math instruction. This includes teachers, instructional coaches, and school administrators. Program is run at a school level, so we will be looking at school level changes. As stated by Hall and Hord (2020) the day-to-day support for implementing change takes place inside the school. The staff of a school share in what the whole school is about, and especially when it comes to implementing change. School leadership, each teacher's expertise, and past experiences with change processes all come together to support and/or inhibit implementing the next change. ● What are the eligibility criteria for applying and participating in the program? o School teams from across the state are eligible to apply for the programs. ● How will you select participants? (For example, random selection, stratified sampling, propensity score matching)? o Using propensity score matching, all PERF-funded schools implementing the NUMeROS and Focus on Algebra programs will be matched with similar schools that are not implementing the selected programs for the purpose of the evaluation. The comparison group will be generated as a 1:1 matched comparison sample using schools that did not implement the program but are otherwise similar. The method for matching will follow What Works Clearinghouse guidelines for appropriate matching characteristics and determinations of baseline equivalence. o For the HQIM program, PED will oversee the selection of participating LEAs through an application process, with priority given to schools serving high percentages of historically underserved students, including Native American students, English learners, students with disabilities, and students

	experiencing poverty. In addition to those student groups, the HQIM program will have a focus on selecting rural schools. ● Additional Information on Funding Allocation and Expenditure ○ Schools are expected to use funds in alignment with program purposes, including professional development, instructional materials, and implementation supports through curriculum based professional learning. Funded schools also agree to monthly progress monitoring tied to performance metrics to ensure implementation is progressing and interventions can be instituted quickly to get applications back on track.

Evaluation Approach and Methods	● Describe how your evaluation will be conducted.
	● Specify what methods and statistical tests PED plans to use.
	  ○ The evaluation will use a matched comparison design to examine the outcomes of participating schools implementing the NUMeROS and Focus on Algebra programs with similar schools that did not receive PERF funding to implement these programs. The evaluation will measure differences between funded schools and a matched comparison group of schools on school-level changes in growth on student achievement in math. For districts selected for the HQIM program, the evaluation will attempt to match districts with other districts that applied for the program but were not selected. If appropriate matches cannot be made between districts that were selected and districts that applied and were not selected, other districts may be selected for the comparison group using similar methods as above, substituting districts for schools.
	  ○ Analyses will include linear regressions and mixed effects models to compare key outcomes from the PERF funded schools and the matched comparison group.
	  ○ Two additional outcomes will be measured using descriptive (measures of central tendency) and/or inferential (t-tests) analyses, without a matched comparison group. An intended short-term outcome of the program is that educators' knowledge of key math knowledge and instructional practices improves. Vendors implementing the programs will observe classrooms in participating schools, providing implementation scores that will be analyzed descriptively and, when possible, inferentially to document changes in classroom practices. Further, the evaluation will provide descriptive and, when appropriate, inferential analysis to describe and document the number of students passing math courses to understand if participating schools experience changes in this metric compared to previous years.
	  ○ To evaluate HQIM implementation with rigor, several instruments and measures will be used. All funded LEAs are required to complete and update the Instructional Materials Implementation Tool. This instrument will allow the evaluation to measure implementation fidelity and progress over time. Schools are also

	required to submit data through the Curriculum Implementation Change Framework, which provides structured monitoring of classroom use of HQIM. The evaluation will test whether higher implementation scores on the Instructional Materials Implementation Tool and the Curriculum Implementation Change Framework are associated with gains in math achievement at the school level. ● If you are planning to assess the program's causal impact on student achievement, explain exactly how this will be done (for example, randomized control trials, matched comparisons). o This quasi-experimental design will assess the program's causal impact on student outcomes from grade levels in participating schools with the same grade levels in matched comparison schools. Given the nature of the funding there may be limitations in the sample size to reach sufficient statistical power and, therefore, affect the reliability and generalizability of the findings. In that case, alternatives will be explored to evaluate the outcomes using descriptive and/or inferential analyses. ● If you cannot assess causal impact: Provide a clear rationale for your proposed evaluation design. o The intended student outcomes of the program are being assessed in a causal design if there is a sufficient sample size of students. However, the evaluation will not be able to assess causal impact of changes in classroom practices or students passing math courses because it would be burdensome on comparison schools to conduct the observations and collect the course completion data that would allow causal comparisons of these outcomes.
Data Collection	● What data collection methods will be used? Who will be responsible for collecting the data? o The evaluation will primarily rely on summative assessment data that is already collected by PED. o Classroom observations will be conducted by the vendors implementing the programs and shared with both PED and the evaluation team. o For grantees with HQIM implementation grants, they will also submit data through the Curriculum Implementation Change Framework that will be collected by the PED and vendors. o Schools that receive funding will work with PED to provide student-level course completion data.
Implementation Timeline	Outline a proposed timeline for activities across the appropriation period:
	Year One (2025-2026): ● Finalize evaluation plan (by 9/1/2025) ● Identify participating schools (by 10/31/2025)

	• Identify matched comparison schools (by 1/31/2026) • Gather data (ongoing- 6/30/2026) • Analyze data (ongoing- 9/1/2026) • Produce progress report with Year 1 outputs and outcomes (by 11/1/2026)
	Year Two (2026-2027): • Review and revise evaluation plan (by 9/1/2026) • Confirm participating schools (by 9/30/2026) • Confirm matched comparison schools (by 9/30/2026) • Gather data (ongoing- 6/30/2027) • Analyze data (ongoing- 9/1/2027) • Produce progress report with Year 2 outputs and outcomes (by 11/1/2027)
	Year Three (2027-2028): • Review and revise evaluation plan (by 9/1/2027) • Confirm participating schools (by 9/30/2027) • Confirm matched comparison schools (by 9/30/2027) • Gather data (ongoing- 6/30/2028) • Analyze data (ongoing- 9/1/2028) • Produce final report comparing key outcomes of funded schools with matched comparison schools over the three-year funding period (by 11/1/2028)
Contacts for Annual Progress Updates	Please use this space to provide primary points of contact at PED for progress updates, expected by November 1 of each year. • Dr. Jacqueline Costales, Ed.D., Executive Manager, Office of Curriculum & Instruction,Jacqueline.Costales@ped.nm.gov • Shafiq Chaudhary, Senior Manager, Math & Science Bureau, Shafiq.Chaudhary@ped.nm.gov

Topic	Math Achievement	Resolution
Reviewers	Tim Bedeaux and Jessica Hathaway, LESC Rachel Mabe and Sarah Dinces, LFC Lori Sciacca and Andrew Miner, DFA	
Plan Strengths	***Program structure and goals:*** ● The submission clearly identifies and describes three distinct initiatives that will be supported by the PERF funding (Focus on Algebra, NUMeROS, and HQIM Implementation). Each has a focus area and intended outcome. ● The problem statement is succinct and grounded in student data. ● Goals are specific to each initiative (improve instructional practice, raise pass rates, and increase student academic achievement). ***Logic model:*** ● All required actor groups (PED, LEAs, educators, and vendors) are identified, and each has detailed activities and outputs assigned. ● Short- and long-term outcomes are explicitly stated, including classroom practice improvements and student performance changes.	

	Evaluation design: ● Propensity score matching to compare participating schools with similar non-participating schools is a valid method for causal inference when randomization is not feasible. This allows for rigor absent causal design, as required by SB201. ● The use of linear regression and mixed effects models are well-suited for analyzing hierarchical education data ● Causal claims are limited to student achievement; instructional practices and course pass rates are analyzed descriptively or inferentially without overreaching claims. ● Combines assessment data, classroom observations, surveys, and course completion data to strengthen validity and triangulate findings. ● The proposed evaluation methods follow What Works Clearinghouse best practices for establishing baseline equivalence and determining credible causal inference in non-randomized designs. ● The plan includes measures of educator engagement and classroom-level implementation, which supports	

	understanding of how and why outcomes occur.	
Requested Changes	***Provide information about how the appropriation will be allocated and expended:*** ● The plan identifies PED as the entity responsible for selecting participating LEAs and awarding HQIM funding, but does not provide detail on the funding process such as how funding amounts will be determined, how recipients will be selected (beyond basic eligibility), or how funding will be distributed over time. ● The plan would benefit from including a description of the funding method (competitive, formula-based, etc.), criteria for LEA selection or prioritization, and a clear timeline (and conditions) for funding disbursement. ● A detailed breakdown of how schools should be expected to use funding would also provide strong context. What is the expected cost of tutoring, professional development, stipends for staff, etc.? ***Within the Underlying Assumptions and Research base section, clarify the evidence base and assign an evidence level:***	Added text to the Target Population and Sampling section providing information on: ● Selection criteria for participating LEAs ● Use of funding Added text to the Underlying Assumptions and Research Base section clarifying the evidence base:

	● Clearly identify the evidence level of each program using definitions provided in the PERF instructions, such as “promising” or “research-based.” PED staff could also add specific findings from internal evaluations or external studies and provide external links where available.	● Added clear identification of evidence level using PERF definitions (e.g., “promising”) ● Added link to previous evaluation
	***Within the target population and sampling section, expand on how selection will occur*:** ● Describe how participating schools will be selected to prioritize equity, including consideration of student demographics, geographic distribution, and academic need. ● This may include a description of which students and educators stand to benefit from the program.	Added text to the Target Population and Sampling section expanding on how selection will occur: ● Selection criteria for participating LEAs, including target student groups and rural schools
	***In the overarching goal section and the evaluation approach and methods section, define thresholds for outcome metrics*:** ● Establishing specific thresholds for success related to key outcomes would offer a greater sense of if the program is on track as monitoring occurs. ● If you have an idea of the desired effect size, this will help accomplish a “power analysis,”	Added text to the Overarching Goals section: ● Neutral language (no specific thresholds until baseline data is analyzed) suggesting that thresholds for each key outcome will be determined once baseline data is reviewed

	giving you more information about how large a sample size you will need to measure a significant effect. ***In the evaluation approach and methods section, strengthen* the *evaluation design for HQIM implementation:*** ● Please provide a more detailed description of how HQIM outcomes will be measured, including specific instruments (for example, the Curriculum Implementation Change Framework) and how this intends to impact student achievement.	Added text to the Evaluation Approach and Methods section providing more detail on HQIM implementation evaluation design: ● Instructional Materials Implementation Tool for measuring implementation fidelity ● Curriculum Implementation Change framework for structured monitoring of classroom use of HQIM
Next Steps	LESC, LFC, and DFA staff request that PED incorporate the above changes and resubmit finalized evaluation plans by the statutory deadline of **September 1, 2025**.	