NEW MEXICO LEGISLATIVE FINANCE COMMITTEE

New Mexico Department of Finance and Administration

Presented November 2026

Government Results and Opportunity Update:

Evidence in Governance, Policy and Programming

Dr. Sarah Dinces
Principal Program Evaluator
Legislative Finance Committee

Simon Miller
Performance Analyst
Department of Finance and Administration

Hearing Roadmap

- **Background**
- **Final GRO Evaluation Reports for inaugural GRO projects**
  - Assessing Agency GRO Final Reports
  - Strengths and Challenges in Final Reports
  - Final Report Summaries by Agency for 2024 appropriations and 2-year 2025 appropriations
- **2026 GRO Evaluation Plans**
- **Challenges and Solutions**
- **Appendices**
  - 2026 Final Evaluation Plan Ratings
  - Determining Evaluation Quality

GRO Background

- In 2024, the Legislature created the Government Results and Opportunity Program and Expendable Trust Funds (GRO), and in 2025, the Public Education Reform Fund (PERF)

- **The goal of GRO and PERF is to determine what works before providing recurring funding through:**
  - **Measuring Meaningful Outcomes:** Assess whether programs **achieve their intended results**, not just spending and activities; and
  - **Rigorous Evaluation Design:** Estimate **causal impacts vs business as usual** using credible comparison groups.

- For the GRO pilots funded from FY25-FY27, they received a total of $214.8M over 3 fiscal years and have spent $135.3M as of the end of October.

- For the select GRO pilots funded for two years (FY26-FY27), they received a total of $32.2M and have spent $16.7M as of the end of October.

Year	FY25	FY26	FY27
2024 appropriations (expended)	$71.6M ($54.9M)	$71.6M ($66.1M)	$71.6M ($14.3M)
2025 2-year appropriations (expended)	-	$16.1M ($13.8M)	$16.1M ($2.9M)
Total	**$71.6M ($54.9M)**	**$87.7M ($79.9M)**	**$87.7M ($17.2M)**

*spend does not include HED for FY27

GRO Background

- GRO fund created in 2024 (HB196)

- Following executive vetoes of evaluation requirements in 2024 and 2025, DFA and LFC encouraged rigorous evaluation from GRO-recipient agencies, but it remained a non-statutory expectation rather than a legal requirement

- **In 2026, the governor signed HB158, formalizing the evaluation framework in statute**

  - LFC budget guidelines approved by the committee in August recommend keeping annual fiscal year spending from the GRO to under 1 percent of the total general fund recurring budget, and to hold PERF recommendations under 1 percent of the total public school support recurring budget.

  - Analysts are directed to consider rubrics results in funding recommendations with GRO projects.

Timeline of GRO

July 1 of appropriation year
- Initial evaluation plan due from agencies to LFC and DFA.

September 1 of appropriation year
- Revised plan from agency due, if necessary.

2026 appropriations

November 1 each year
- Annual evaluation update due from agency.

September 1 of final appropriation year
- Final evaluation report from agency due.

2024 appropriations

Final Report Submissions

- Reports were assessed using criteria provided to agencies

Goals	Study quality	Data collection	Results of evaluation
Clearly defined goals that were measurable	Causal evaluation with moderate to high rigor	Data collected in a reliable way with adequate sample size	Reliable, statistically significant positive impact
Most elements clearly defined/ measurable	Causal evaluation limited rigor	Data mostly collected reliably or with a close to adequate sample size	Overall mostly positive findings, potentially some null or ns results
Some goals are defined and measurable but others are not	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient N size
Unclear or unmeasurable goals	Implementation and outputs evaluation (non-causal)	Data collected in an unreliable way or with an inadequate sample size	Mostly null results or negative
No goals	No evaluation or anecdotal evidence.	No data collected	Null or negative results

- DFA and LFC team reviewed each report and provided this information to analysts when considering budget recommendations, with funding guidelines on the following slide.

Final Report Submissions

Evaluation results for 25 appropriations from 2024 and select 2025 appropriations were due from 12 agencies on September 1, 2026

- ECECD, ALTSD, HCA, WSD, OFRA, CYFD, 2nd DA, AOC, PED, HED, SFCC, NMIMT
- $247M funded for the inaugural projects, of which $152M was spent as of the end of October.

The reports and ratings will inform the Legislature and Executive about whether to continue funding.

LFC Staff Guidelines	
Outcome	Potential Recommendations
Positive Impact	Fund in base or continue nonrecurring (mostly greens)
Unclear Impact	Continue nonrecurring funding with stipulations (mixed)
Negative/No Impact	Do not fund (mostly yellows/oranges/reds)

* Staff decisions were not made solely on the outcome but on the combination of the other factors and the outcome.

Final Report Submissions: Strengths and Challenges

- **Strengths:**
  - On-time submissions
  - Clear articulation of goals

- **Challenges:**
  - Limited evaluation rigor leading to **limited evidence of impact**
  - Most evaluations (84%; 21 of the 25 appropriations) were limited to reporting program activities, outputs, and outcomes and did not support generalization, **leading to uncertain results**
  - Few (16%; 4 out of 25) attempted to determine whether the programs themselves *caused* improvements, even though the intent was to understand cause

- Some of the 2025 two-year appropriations were not well suited for impact evaluation

- **Main Conclusion: More rigorous evaluation design or study quality is needed.**

Final Report Tentative Recommendation Summary

Preliminary LFC Staff Recommendations		
Potential Recommendation	Number of Appropriations	Total GRO Funding Amount
Fund in base or continue recurring funding	5	$14.5M/year
Continue nonrecurring funding with stipulations	9	$51M/year
Do not fund	11	$17.6M/year

2024 Appropriations

ECECD-Wage and Career Ladder

- ECECD established a ladder to determine salaries and provide monthly wage supplements for childcare educators and administrators with the goal of improving recruitment and retention.

- **The evaluation examined outcomes for the wage and career ladder, showing some positive results, but did not establish causality leading to uncertainty.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Wage and Career Ladder	$15,000.0	$13,710.9	Clearly defined goals that were measurable	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Overall mostly positive findings, potentially some null or nonsignificant results

ALTSD-New Mexicare

- ALTSD established New Mexicare to provide unpaid caregivers a stipend as well as training, and support services for providing care to elderly individuals living in a private residence but do not qualify for Medicaid caregiving services.

- **ALTSD used a comparison group and found mixed effects, and there were questions about the appropriateness of the comparison group**.

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/2 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
New Mexicare	$9,375.0	$4,066.2	Most elements clearly defined/ measurable	Causal evaluation limited rigor	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

HCA – SBIRT and CCBHCs

- HCA used GRO funds to increase the capacity to provide Screening, brief intervention, and referral to treatment (SBIRT) and establish certified community behavioral health clinics (CCBHCs).
  - SBIRT training can help healthcare providers learn to screen patients, provide brief interventions, and refer individuals needing additional care to appropriate treatment services.
  - CCBHCs provide coordinated behavioral health services. Aiming to improve timely access to care, particularly for underserved populations.

- **HCA did not evaluate the impact of these funds, leading to uncertainty about the outcomes of these evidence-based practices in New Mexico.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
SBIRT and CCBHC expansion	$15,000.0	$10,200.9	Most elements clearly defined/ measurable	Implementation and outputs evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

WSD- BeProBeProud and Preapprenticeships

- BeProBeProud uses a mobile workshop with simulators and activities to introduce students to trades and technical careers, aiming to increase awareness and connect interested students
- Pre-apprenticeships provide youth with up to 400 hours of paid work experience, supplemented by coaching and employment preparation.
- **Neither appropriation assessed causal impacts; without causal evaluation and strong data collection, the results were unclear for both appropriations.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Trades Career Exploration (BeProBeProud)	$6,000.0	$4,854.7	Most elements clearly defined/ measurable	Implementation and outputs evaluation (non-causal)	Data collected in an unreliable way or with an inadequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Pre-apprenticeships	$1,800.0	$1,200.0	Most elements clearly defined/ measurable	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

OFRA- Multidisciplinary Teams

- OFRA's Interdisciplinary Legal Team program provides support to parents involved in child welfare through a team approach. The program aims to help parents address barriers to case resolution, reduce the time children spend in foster care, and achieve safe, permanent outcomes.

- **The evaluation suggests some promising outcomes regarding length of time to permanency, but since it was non-causal, the impacts are uncertain.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Multidisciplinary Team Legal Services	$4,875.0	$3,560.4	Most elements clearly defined/ measurable	Implementation and outcomes evaluation (non-causal)	Data mostly collected reliably or with a close to adequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

Children, Youth and Families

- CYFD received 4 GRO appropriations focused on preventing child maltreatment and strengthening the foster care system.
- 2 of the 4 appropriations did not use causal evaluation so the impact was uncertain for these appropriations. MLR attempted to measure causality but lacked an adequate sample size and didn't assess defined outcomes.

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Recruitment of social worker case workers	$1,687.5	$1,088.8	Most elements clearly defined/ measurable	Implementation and outputs evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Multilevel response (MLR)	$4,200.0	$1,249.7	Most elements clearly defined/ measurable	Causal evaluation limited rigor	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Treatment foster	$3,750.0	$2,216.3	Clearly defined goals that were measurable	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Prevention and Intervention Program Expansion	$9,000.0	$2,601.2	Most elements clearly defined/ measurable	Causal evaluation limited rigor	Data collected in an unreliable way or with an inadequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

Public Education

- For the educator clinical practice appropriation, PED reported some **promising outcome data**, but since the comparison groups were of varying quality, **leading to uncertainty about the program's impact**. Further, more time is needed to collect data to assess residents' impact on student achievement as independent teachers.
- For the special education stipends appropriation, PED reported promising year one retention outcomes, but **program implementation changed significantly in year two (FY26), limiting comparability across years**.

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Clinical Practice (Teacher Residency Program)	$60,000.0	$43,481.2	Clearly defined goals that were measurable	Implementation and outcomes evaluation (non-causal)	Data mostly collected reliably or with a close to adequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Special Education Stipends	$15,000.0	$8,603.2	Most elements clearly defined/ measurable	Implementation and outcomes evaluation (non-causal)	Data mostly collected reliably or with a close to adequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

Higher Education

- The workforce certificate appropriation funds short-term workforce training at two-year institutions, as well as related internships and apprenticeships, to help students earn credentials, improve employment, and meet workforce needs.
- Integrated Education and Training (IET), a national best practice, combines adult education and ancillary services with workforce preparation and occupational training.
- **While evaluations examined outcomes for both appropriations, it did not establish causality, leading to uncertain results. Further study is needed to understand the impacts of these appropriations.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 7/1/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Noncredit workforce certificates	$60,000.0	$33,083.4	Most elements clearly defined/ measurable	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Integrated education and training	$6,000.0	$3,405.6	Clearly defined goals that were measurable	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

New Mexico Institute of Mining and Technology

- NMIMT developed the Geothermal Research and Education Center to expand knowledge of geothermal resources and support their potential development and use.
- **The evaluation provides a descriptive assessment of the center's activities and accomplishments but does not identify or measure outcomes.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Geothermal	$1,500.0	Not in SHARE	Some goals are defined and measurable but others are not	Implementation and outputs evaluation (non-causal)	Data mostly collected reliably or with a close to adequate sample size	Unclear or mixed results likely due to evaluation design or insufficient N size

Santa Fe Community College

- SFCC received funding to provide suicide prevention training. DOH also receives funding to provide suicide prevention training. SFCC developed its own model for suicide prevention training.

- **SFCC provided promising short-term outcome data (including participants' perceptions of their abilities); it did not assess causality or use other trainings shown to reduce suicide as a benchmark, leading to uncertainty about the impact.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Suicide prevention	$1,500.0	Not in SHARE	Some goals are defined and measurable but others are not	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

2025 2-Year Appropriations

Administrative Office of the Courts

- Pretrial services supervise defendants released pending trial, monitor compliance with court orders, provide hearing reminders, identify potential violations, and prepare investigations and risk assessments to inform judicial decisions. The program aims to improve appearance rates, promote public safety, and improve compliance.
- The New Mexico Courts HelpLine assists individuals navigate the state court system. Staff provide information about court procedures, forms, and resources; help identify appropriate court services; and refer callers to organizations.
- **Neither appropriation looked at cause, leading to uncertainty about the impact of the appropriations.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Pretrial Services	$2,400.0	$1,109.1	Some goals are defined and measurable but others are not	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Turquoise call center	$1,600.0	$800.0	Some goals are defined and measurable but others are not	Implementation and outputs evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

2nd Judicial District Attorney

- 2nd DA received a GRO appropriation for court monitors and to implement changes to New Mexico's criminal competency procedures under Chapter 4, Laws 2025 (HB 8). Given the administrative nature of the activities funded, the initiative is not particularly well suited to the GRO fund's emphasis on causal evaluation and measurable program outcomes.

- **The evaluation did not measure outcomes, leading to uncertainty about the impact of the appropriations.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Court Monitors	$500.0	$331.3	Most elements clearly defined/ measurable	Implementation and outputs evaluation (non-causal)	Data mostly collected reliably or with a close to adequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

Children, Youth and Families

- CYFD received 6 2-year 2025 appropriations to cover a wide range of topics, mostly related to the *Kevin S.* settlement.

- **While many of these appropriations did not lend themselves to a causal evaluation, limited outcome data was provided and there was no assessment of causality, leading to uncertainty related to outcomes.**

Short Title	Total $ ($1,000s)	Total $ Spent as of 10/31/26 ($1,000s)	Goals	Study Quality	Data Collection	Outcomes
Personnel re office child Advocate	$600.0	$55.0	Unclear or unmeasurable goals	Implementation and outputs evaluation (non-causal)	Data mostly collected reliably or with a close to adequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Case aides	$9,600.0	$6,086.6	Most elements clearly defined/ measurable	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Regional emergency response teams	$5,600.0	$2,146.1	Unclear or unmeasurable goals	Implementation and outputs evaluation (non-causal)	No data collected	Unclear or mixed results likely due to evaluation design or insufficient sample size
Remedial order personnel	$1,800.0	$1,135.9	Most elements clearly defined/ measurable	Implementation and outputs evaluation (non-causal)	Data collected in an unreliable way or with an inadequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Foster care maintenance payment increase	$5,000.0	$2,500.0	Clearly defined goals that were measurable	Implementation and outcomes evaluation (non-causal)	Data collected in an uncertain way or with an uncertain sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size
Operating expenses	$5,200.0	$2,483.9	Some goals are defined and measurable but others are not	Implementation and outputs evaluation (non-causal)	Data collected in an unreliable way or with an inadequate sample size	Unclear or mixed results likely due to evaluation design or insufficient sample size

2026 Evaluation Plans Reports

- A necessary component of strong evaluation plans is to ensure that they allow the agency and others to understand whether the appropriation ***caused*** a change in outcomes.

- Of the 8 appropriations with final plans this year:
  - 4 of the 8 should allow for an assessment of cause.
  - 2 of the 8 should allow for a measurement of outcomes without knowing cause
  - 2 of the 8 did not provide a plan, but both stated they will contract for an evaluation

- Additionally, all appropriations have clear logic models, most have strong performance measures, and all plan to have final reports released publicly or to DFA/LFC.

Challenges and Solutions with GRO

Challenges	Solutions
Original legislation did not include evaluation requirements, and initial final reports lacked causal analysis	The Legislature passed and the Governor signed HB158 in 2026
Uncertainty regarding what agencies are doing over the course of the funding	DFA/LFC workgroup requiring November updates from agencies to highlight changes in implementation that need to be discussed
GRO funding is requested and used for administrative tasks that don't fit within an impact evaluation framework	GRO fund should be reserved for piloting projects that fit within an impact evaluation framework
Low-quality evaluation	DFA/LFC workgroup in discussions to increase technical assistance capacity

Questions?

Sarah Dinces
Sarah.dinces@nmlegis.gov
(505) 986-4571

Matt Goodlaw
Matt.Goodlaw@nmlegis.gov
(505) 986-4515

Simon Miller
Simon.miller@dfa.nm.gov
(505) 479-2963

Shanthi Bangi
Shanthi.Bangi@dfa.nm.gov
(505) 394-5286

Appendix: 2026 GRO Appropriation Final Plan Ratings

2026 Final Evaluation Plans

Agency	Short Purpose	Years of Funding	Logic Model	Performance Measure	Evaluation Plan	Communication of Results
AOC	For the expansion of assisted outpatient treatment programs, a competency diversion pilot and other behavioral health programs, for expenditure in fiscal years 2027 and 2028	2	Most elements clearly defined.	Most are clearly defined, mixture of outputs and outcomes.	Causal evaluation with Low Rigor	Public report
LOPD	To pilot hourly rates for contract attorneys	3	Most elements clearly defined.	Most are clearly defined, mixture of outputs and outcomes.	Implementation and outcomes evaluation (non-causal)	Public report
DFA	For distributions to state agencies pursuant to a rate schedule adopted pursuant to Section 15-3B-18 NMSA 1978.	3	Most elements clearly defined.	Most are clearly defined, mixture of outputs and outcomes.	Implementation and outcomes evaluation (non-causal)	Report to DFA/LFC
ECECD	To continue implementation of the wage and career ladder framework.	3	All elements clearly defined.	Most are clearly defined, mixture of outputs and outcomes.	Causal evaluation with Low Rigor	Public report

2026 Final Evaluation Plans

Agency	Short Purpose	Years of Funding	Logic Model	Performance Measure	Evaluation Plan	Communication of Results
WSD	For the office of housing operations	3	Most elements clearly defined.	Most are clearly defined, mixture of outputs and outcomes.	Implementation and outcomes evaluation (non-causal)*	Public report
CYFD	For startup and operational costs of evidence-based programs delivered in a community-based setting, including youth mentoring services, for high-risk youth within the juvenile justice facilities program	3	Most elements clearly defined.	Performance measures written, but not clearly defined.	Implementation and outcomes evaluation (non-causal)*	Public Report
CYFD	For a dedicated SafeCare unit within the protective services program to provide in-home services to families reported to the agency for suspected abuse or neglect	3	All elements clearly defined.	Clearly defined, mixture of outputs and outcomes	Causal evaluation limited rigor	Public report
DOT	To perform road safety audits and site assessments on state and local roads	3	All elements clearly defined.	Clearly defined, mixture of outputs and outcomes.	Causal evaluation with moderate to high rigor	Public report

*WSD and CYFD stated they would contract for an evaluation examining outcomes but did not provide details. CYFD stated it would be a causal evaluation.

Appendix:
Determining Research Quality

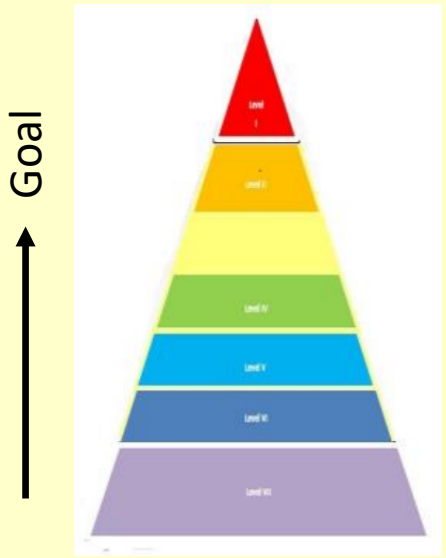
Evidence Pyramid

Research quality also depends on:

1. Soundness of the logic model
2. Quality of the study
3. Independence of the study
4. Adequacy of statistical tests
5. Adequacy of the reporting

The aim is to see evaluations that are moving toward more rigor.

Evidence Pyramid.



Level 1: Systematic Reviews & Meta-analysis of RCTs; Evidence-based Clinical Practice Guidelines

Level 2: One or more RCTs

Level 3: Controlled Trials (no randomization)

Level 4: Case-control or Cohort study

Level 5: Systematic Review of Descriptive and Qualitative studies

Level 6: Single Descriptive or Qualitative Study

Level 7: Expert Opinion

Source: UNM