Shared Responsibility, Stronger Evidence:
A Co-Designed Partnership Framework for Trust & Accountability in Education

Joshua Stewart, Ph.D., Rocky Mountain Research & Strategy • Yi Chieh Newton, Ph.D., Wilson Language Training®

Building credible evidence starts with the question and context.

We co-design studies that are feasible, rigorous, and useful to real-school settings because they reflect real-world implementations.

Partnership Contributions

School Systems:
Priorities, context, implementation reality, local data, policy environment

Curriculum Developers:
Program curriculum design rationale, implementation supports, operational/usage data, cross-site knowledge

Independent Researchers and Evaluators:
Methodological and analytic expertise, scientific rigor, independent judgment, and objective interpretation

Co-Designed Evidence-Building Framework

This iterative process adapts to context and accumulates learning over time.

1 Shared Evidence Opportunity
What do we jointly need to know?

2 Readiness Assessment
What can this setting support?

3 Evidence Purpose
What research questions matter?

4 Design Decision
What is the strongest defensible design?

5 Undertake, Interpret & Learn
What did we learn and what remains uncertain?

Readiness Domains

- Partnership Strength
- Implementation Capacity
- Data Infrastructure
- Setting Context
- Design Feasibility

Design Options:

- Randomized Trial (RCT)
- Quasi-Experimental (QED)
- Regression Discontinuity (RDD)
- Interrupted Time Series (ITS)
- Single-Case Experimental (SCED)
- Hybrid Effectiveness–Implementation (Type 1, 2, or 3)
- Descriptive/Implementation Study

If Conditions Change:

- Reassess readiness
- Adapt the design
- Narrow claims
- Preserve useful learning

Next Evidence Need

Key Design/Input Considerations

- Assignment process/ Comparison availability
- Baseline data/Pre-trends
- Outcome quality and alignment
- Sample size and power
- Implementation fidelity and variation
- Timing and policy changes
- Data linkage and quality
- Ethical and practical constraints

Example: Anonymized Case Evidence Assessment

Strong evidence · Some evidence · Limited evidence

Evidence Domain	Program A	Program B
Effectiveness	Strong	Limited
Implementation	Some	Limited
Context & Variation	Some	Some
Mechanisms	Limited	Limited
Feasibility & Usability	Strong	Some

Identified gap (e.g., implementation in Program B) → Next study designed to fill that gap

Building a Cumulative Evidence Portfolio

Different studies answer different questions. Together they create a complete, credible picture.

Effectiveness
Does it improve intended outcomes?
- Student outcomes
- Educator outcomes
- Subgroup effects
- Replication across time

Implementation
How is it implemented, and how does implementation relate to outcomes?
- Adoption, dosage, fidelity
- Reach and penetration
- Implementation/outcome links
- Sustainability

Context & Variation
For whom, where, and under what conditions?
- School/leader conditions
- Student population differences
- Treatment effect heterogeneity
- Adaptation and local fit

Mechanisms
What components and processes drive outcomes?
- Active ingredients
- Mediators and pathways
- Coaching/support processes
- Why and how it works

Feasability & Usability
Can it be used, adapted, and sustained in real settings?
- Acceptability and burden
- Practicality and fit
- Costs and resources
- Usability and tool readiness

Why This Matters

School systems, curriculum developers, independent researchers, and evaluators each face constraints that make traditional evaluation roadmaps hard to follow.

Rigor ↔ Feasibility:
High standards must work within real-world conditions.

Consistency ↔ Adaptation:
Implementation varies across schools, leaders, and students.

Timely Learning ↔ Certainty:
Decisions can't wait for the "perfect" study.

Credibility Commitments

Transparent
Methods, data, and limitations are made clear.

Independent Judgment
Research decisions are guided by evidence and ethics.

Contestable
Findings can be questioned, tested, and improved.

Proportional Claims
Claims match the strength and scope of the evidence.

Local Relevance
Evidence is meaningful and usable for decision-makers.

AM_RMRS_Poster_V4.indd 1 1/9/2026 10:19:am