Center for Education Policy Research
HARVARD UNIVERSITY

STRATEGIC DATA PROJECT
2026 CONVENING

PROOF POSITIVE

Writing Actionable Insights: Turning Evidence into Clear, Usable Messages

Wednesday, May 13th 1:15-2:30PM

EdResearch for Action

Annenberg Institute
BROWN UNIVERSITY

Writing Actionable Insights

How to Translate Education Research for the People Who Need It Most

Christina Claiborne
Annenberg Institute at Brown University

Carrie Conaway
Harvard University

The translation problem in education research is not a research problem.

It is a writing problem.

And writing problems can be solved.

Who We Are

Christina
Claiborne

Research Project Director

Annenberg Institute at Brown University

Carrie
Conaway

Senior Lecturer on Education

Harvard Graduate School of Education

KEY OBJECTIVE

Build skills to make research findings actionable

By the end of this session, you will be able to:

01

Diagnose

Identify what makes research writing hard for practitioners to use — and where your own drafts get in the way.

02

Translate

Apply six concrete principles to turn dense findings into prose that lands with non-academic audiences.

03

Revise

Practice the moves on real examples so the techniques become habits, not one-time fixes.

What you write shapes what gets implemented

The education research community has built an enormous body of knowledge about what works and what doesn't, and then packaged it in a form that almost no practitioner will use. When research is hard to read, the decisions get made anyway — just with less of your evidence in them.

Writing for action is how your work gets used

Six principles for writing actionable insights

01 The "so what" test

02 The most striking sentence

03 Rigor vs. clarity

04 Mixed evidence — and what to do with it

05 Making numbers memorable

06 Translating regression results

01

PART ONE

The "So What" Test

What may be obvious to you is rarely obvious to your audience.
Make the implication explicit.

The "so what" is whatever the reader should do or understand differently

1. **What's the finding?**

   EXAMPLE: "Districts are currently struggling to staff special education positions, for both special education teachers and paraeducators."

2. **What's the implication?**

   EXAMPLE: Vacancies don't mean classrooms are empty. They mean those classrooms are filled with under-credentialed staff, long-term substitutes, or general education teachers covering caseloads they weren't trained for

3. **Why should I care?**

   EXAMPLE: More than half of school districts are struggling to hire qualified special education teachers — and over 800,000 students are being taught by under-qualified personnel

> ***What changed:*** *the before is about a district problem. The after is about students being harmed — and quantifies the scope. District leaders care about the issue when they see who's affected and how many.*

Try it!

Try it #1

1

What's the finding?

EXAMPLE: "Researchers, policymakers, and educators have debated the extent to which ML-ELs' primary languages should be used in school despite evidence that bilingual and dual language instruction benefits students. English-only models are the norm."

2

What's the implication?

3

Why should I care?

Try it!

Try it #1

1. **What's the finding?**
   EXAMPLE: "Researchers, policymakers, and educators have debated the extent to which ML-ELs' primary languages should be used in school despite evidence that bilingual and dual language instruction benefits students. English-only models are the norm."

2. **What's the implication?**
   EXAMPLE: Despite the evidence, most schools default to English-only instruction, meaning the approach with the strongest research base is the one ML-EL students are least likely to receive.

3. **Why should I care?**
   EXAMPLE: The vast majority of ML-EL students are not being educated through bilingual and dual language models, despite evidence that these programs are more effective than English-only instruction.

> ***What changed:*** *the after names the gap between what the research shows and what students experience. That gap is the engine of the "so what."*

02

PART TWO

The Most Striking Sentence

"The brain craves meaning before detail… If you want people to pay attention, don't start with details. Start with the key idea and, in a hierarchical fashion, provide details that support the big picture."

— brain researcher and professor John Medina

The most striking sentence is likely already in your draft.

The exercise: Read your own draft and find the single sentence a school board member, principal, or policymaker would find **most striking**. This is the one that, if they read nothing else, would make them stop and pay attention. That sentence belongs at the top.

Three questions to surface it

1 Does it name a person, place, or number?

Concrete beats abstract. Look for proper nouns, percentages, ratios, dollar amounts.

2 Does it name the stakes?

Practitioners read past institutional problems and stop on student-impact problems. Ask: who specifically is worse off because of this finding?

3 Does it expose a gap between what we know and what we do?

Interest often comes from juxtaposition: we know X works — most students aren't getting X. Look for places the piece documents both a finding and a failure to act on it.

Organize your structure around the reader's decision, not your research process

If your most striking fact is buried on page 4, your structure is wrong.

TYPICAL OPENING

"This brief examines patterns in teacher retention across U.S. public schools, drawing on national survey data and state administrative records to identify factors associated with attrition…"

LEAD WITH THE MOST INTERESTING THING

"Half of new teachers leave within five years — and the schools serving the children with the greatest needs lose them fastest."

03

PART THREE

The Researcher's Dilemma: Rigor vs. Clarity

The academic bar and the practitioner need are genuinely different. A finding that would be considered preliminary in a journal might be well-established enough to act on in a school.

Rigor looks different when your reader is a superintendent than when your reader is a peer reviewer.

The peer-review reflex	*Write as if every sentence will be scrutinized by someone who wants to find you wrong*	→	State what the weight of evidence supports, clearly and directly
The correlation-causation fear	*Avoid implying causation even when practitioners need to act*	→	Name the limitation once; don't let it swallow the finding
The limitations section	*Listing every way the study could be wrong*	→	Surface only the limitations that affect what practitioners should do.
The 'more research needed' habit	*Signal that the question isn't settled*	→	State what the majority of evidence supports and what judgment and local context should fill the gaps.

Find the accurate middle: the statement that is as confident as the evidence warrants.

Defensible to peers. Unreadable to anyone else. The people who most need the research cannot use it at all.

Honest about what you found. Accessible to the decision maker.

Reads cleanly but states findings with more confidence than the evidence warrants

This calibration is not a formula. It is a judgment. Like all judgments, it improves with practice, with feedback, and with genuine attention to what it means to be honest about a body of evidence that is almost always more complex than a single clean conclusion.

Give readers the actual level of confidence the evidence supports — in language they can act on.

USE THIS PHRASE	WHEN THE EVIDENCE IS…
"The evidence is clear" *"Research consistently shows"*	**STRONGEST** Multiple rigorous studies point in the same direction.
"Evidence suggests" *"Research indicates"*	**STRONG** Solid evidence with some variation. Use for the large middle of the literature.
"Early evidence suggests" *"Preliminary findings point to"*	**EMERGING** Promising but limited evidence. Use for new interventions or emerging areas.
"Researchers hypothesize" *"This is likely because"*	**INDIRECT** Theoretical reasoning or indirect evidence. Use when explaining mechanisms not directly tested.
"The evidence is mixed" *+ explanation of the pattern*	**CONTESTED** Genuine inconsistency. Never a full stop — always explain the sources of variation and what the mixed pattern means.

Try it #2

Try it! Rewrite the excerpt using the calibration vocabulary

Identify the right tier for the finding, then rewrite the excerpt so a practitioner can act on it.

BEFORE

Initial pilot evaluations of teacher residency programs indicate that participants may exhibit higher rates of three-year retention than traditionally-prepared peers, though the available evidence is largely drawn from a small number of programs in urban districts and further research is warranted.

YOUR REWRITE

Try it #2

Try it!

BEFORE

Initial pilot evaluations of teacher residency programs indicate that participants may exhibit higher rates of three-year retention than traditionally-prepared peers, though the available evidence is largely drawn from a small number of programs in urban districts and further research is warranted.

AFTER

Early evidence suggests that teacher residency programs improve three-year retention. The findings come from a small number of urban districts, so the pattern needs to be tested elsewhere before scaling.

The move: *The original uses "initial," "pilot," "may," "indicate," and "further research is warranted" — five hedges for one finding. The rewrite uses one calibration phrase, then names the limit honestly in plain language.*

04

PART FOUR

Mixed Evidence — and What to Do With It

Decision-makers can't wait for more research; they're going to act anyway.

"The evidence is mixed" is a setup, not a closer

THE FULL STOP

What writers usually do

"The evidence is mixed."

Treats inconsistency as the end of the sentence.

Tells the reader nothing about what the mixed pattern means — or what they should do with it.

THE SETUP

What to do instead

"The evidence is mixed — and the pattern is informative."

Treats inconsistency as a finding to interpret. The next sentence explains what the mixed pattern reveals.

Mixed findings often have a structure. Naming it gives the reader something to act on.

Four patterns that can be hiding inside "mixed evidence"

1 WORKS FOR SOME, NOT OTHERS

The intervention helps a specific subgroup (by age, demographic, school context, or starting level) but not others.

2 EFFECTS DEPEND ON DOSAGE

The intervention works above a threshold of intensity or frequency and produces little or nothing below it.

3 EFFECTS DEPEND ON WHAT ELSE IS IN PLACE

The intervention works when paired with complementary supports and fails in isolation.

4 THE MAJORITY POINTS ONE WAY

The majority of studies agree on the direction of the effect; a few studies create the appearance of mixedness. The honest summary names the majority.

THE MOVE · *When you encounter mixed evidence, ask if any of these four patterns fit. The pattern is the story.*

Phrasing the pattern, not the inconsistency

Each pattern has a sentence shape that does the interpretive work for the reader.

PATTERN	PHRASING THAT WORKS
WORKS FOR SOME, NOT OTHERS	**"The evidence is consistent within [subgroup] and weaker elsewhere."** *"The question isn't whether X works — it's for whom it works."*
EFFECTS DEPEND ON DOSAGE	**"The intervention shows clear effects above [threshold] and little below it."** *"Studies of high-intensity programs find consistent gains; lower-intensity versions don't."*
EFFECTS DEPEND ON WHAT ELSE IS IN PLACE	**"X works when paired with [supporting condition]. Without it, X is not effective"**
THE MAJORITY POINTS ONE WAY	**"The weight of the evidence shows [direction]; a smaller body of work disagrees."** *"'Mixed' overstates the disagreement: most rigorous studies point in the same direction."*

Example

BEFORE

"Evidence on the impacts of teacher induction is quite mixed, with several high-profile studies showing little to no effect. However, recent studies indicate that induction can have positive effects, questioning the non-effects of these earlier studies."

AFTER

"The majority of high-quality studies show that induction improves teacher retention, instructional quality, and student learning outcomes."

What changed: *the before opens with "mixed" — and the reader stops processing there. The after leads with the direction the weight of evidence points, and names what induction actually improves.*

Try it!

BEFORE

"The evidence on class-size reduction is mixed. Some studies find positive effects on student achievement, particularly for younger students and disadvantaged populations, while others find smaller or null effects. More research is needed to identify the conditions under which class-size reduction is most effective."

YOUR REWRITE

Try it #3

Try it!

BEFORE

"The evidence on class-size reduction is mixed. Some studies find positive effects on student achievement, particularly for younger students and disadvantaged populations, while others find smaller or null effects. More research is needed to identify the conditions under which class-size reduction is most effective."

AFTER

Class-size reduction works for younger students and disadvantaged populations. The evidence is thinner, and the effects smaller, for everyone else.

The move: *The original sentence already contains the answer, but buries it inside a "mixed evidence" frame that suggests we don't know what we're looking at. We do. The pattern is conditional, not contested. The rewrite just promotes what the original sentence already knew to the main clause.*

Sometimes "inconclusive" really is the answer

Honest mixed evidence is a real category. Don't manufacture a pattern where none exists.

1

NAME IT HONESTLY

Genuinely unresolved is a finding.

Well-designed studies disagree. The field hasn't resolved why. Forcing a tidy interpretation overstates what we know — and erodes the trust your next brief depends on.

2

APPLY THE TEST

Are you describing the evidence, or imposing a story?

If the pattern requires squinting to see, the honest move is:

"Too little is known to recommend with confidence."

05

PART FIVE

Making Numbers Memorable

"Statistics aren't inherently helpful. It is the scale and context that makes them so."

— Chip Heath & Dan Heath, Made to Stick

“We do hard, painstaking work to generate the right numbers to help make a good decision- but **all that work is wasted if these numbers don’t take root in the minds of decision makers**. As lovers of numbers, we find this tragic. There’s no guidebook for the fundamental process of getting people to understand the numbers.”

Chip Heath and Karla Starr
Making Numbers Count

Three moves to make a number stick.

Pick the move that fits the number.

01

Make it COMPARATIVE

Show what it's relative to

Numbers gain meaning from comparison — to averages, to prior years, to benchmarks the reader already understands.

02

Make it FAMILIAR

Anchor in everyday terms

Translate numbers into time, money, and outcomes the reader can picture.

03

Make it HUMAN-SCALE

Translate to people

Convert percentages into the number of students, teachers, or schools actually affected in the reader's context.

The principles of comparative, familiar, and human-scale are adapted from "Made to Stick" by Chip Heath and Dan Heath (2007).

COMPARATIVE

Six types of comparisons

Compare to Similar Others

Show each school's data alongside all others to put findings in context.

Compare to Average / Business as Usual

Benchmark against a familiar reference point.

Compare Over Time

Translate change into percentage terms that resonate.

Compare to Familiar Benchmarks

Anchor in everyday knowledge and human scale.

Compare to Zero

Surface the share of students who never participated at all.

Compare to a Hypothetical

Quantify the gap by imagining parity — how many more students would succeed under it.

COMPARATIVE

Six types of comparisons

Compare to Similar Others

Show each school's data alongside all others to put findings in context.

Seamless college enrollment rates vary substantially across FCS high schools.

Compare to Average / Business as Usual

Benchmark against a familiar reference point.

15% of special ed teachers leave annually vs. 12% for other teachers.

Compare Over Time

Translate change into percentage terms that resonate.

Algebra I enrollment rose 13 ppts — a 50% increase since the policy was introduced.

Compare to Familiar Benchmarks

Anchor in everyday knowledge and human scale.

Among Fortune 500 CEOs, there are fewer women than men named James.

Compare to Zero

Surface the share of students who never participated at all.

38% of economically disadvantaged students took zero AP exams, vs. 11% of non-ED students.

Compare to a Hypothetical

Quantify the gap by imagining parity — how many more students would succeed under it.

If ED students passed AP exams at the same rate as non-ED students, 142 more would pass each year.

COMPARATIVE

BEFORE

Test scores from NAEP released this year show that 33 percent of eighth graders are reading at a level that is "below basic."

AFTER

Test scores from NAEP released this year show that 33 percent of eighth graders are reading at a level that is "below basic" — meaning that they struggle to follow the order of events in a passage or to even summarize its main idea. That is the highest share of students unable to meaningfully read since 1992.

Two anchors: *first, name what "below basic" actually means in plain language. Then anchor 33% to a familiar reference — the worst since 1992 — so the reader feels the size.*

Adapted from: theatlantic.com/ideas/archive/2025/10/education-decline-low-expectations/684526/

FAMILIAR

Anchor data in everyday language or experiences

People relate to time, money, and effort more than raw stats.

BEFORE

"Mentor teachers receive a $200 stipend per semester."

TRANSLATED

> *"Mentor teachers are paid $200 per semester — equivalent to just $4 per hour for 50+ hours of mentoring."*

BEFORE

"Starting teacher salaries average $44,500"

TRANSLATED

> *"After taxes and a typical student loan payment, that's roughly $2,400 a month. This is less than the median rent in most metro areas where these teachers work."*

Try it #4

HUMAN-SCALE

Try it!

Put data in terms of people, not just percentages

Translate percentages into real people

BEFORE

"Only 28% of 8th graders were proficient in math in 2024."

YOUR REWRITE

Scale small effects to system-level impact

BEFORE

"Each additional day of tutoring in a 20-day program is associated with a 0.5 percentile point gain on the math test."

YOUR REWRITE

Try it #4

HUMAN-SCALE

Put data in terms of people, not just percentages

Translate percentages into real people

BEFORE

"Only 28% of 8th graders were proficient in math in 2024."

TRANSLATED

"Only 28% of 8th graders were proficient in math in 2024. In a district of 1,000 students, that's just 280 on grade level."

Scale small effects to system-level impact

BEFORE

"Each additional day of tutoring in a 20-day program is associated with a 0.5 percentile point gain on the math test."

TRANSLATED

"Students who attend the full 20-day program would gain a full point — equivalent to a 10 percentile point jump for an average student."

06

PART SIX

Translating Regression Results

A coefficient describes the data. A predicted value describes a person.

1 Report predicted values from regressions, rather than coefficients

Why it works: Regression coefficients tell researchers how variables relate. But predicted probabilities or counts tell readers what a specific person or school can expect — which is far more actionable.

2 Compare effect sizes to other, more familiar or relatable effects

Why it works: Effect sizes like 0.20 SD mean nothing to most readers. Comparing to a known, relatable effect instantly calibrates magnitude without requiring statistical training.

1 Report predicted values from regressions, rather than coefficients

✗ THE COEFFICIENT WAY

"CTE concentrators earn 3% more annually five years after graduation."

Reader: What's the baseline? Is 3% a lot?

✓ PREDICTED-VALUES WAY

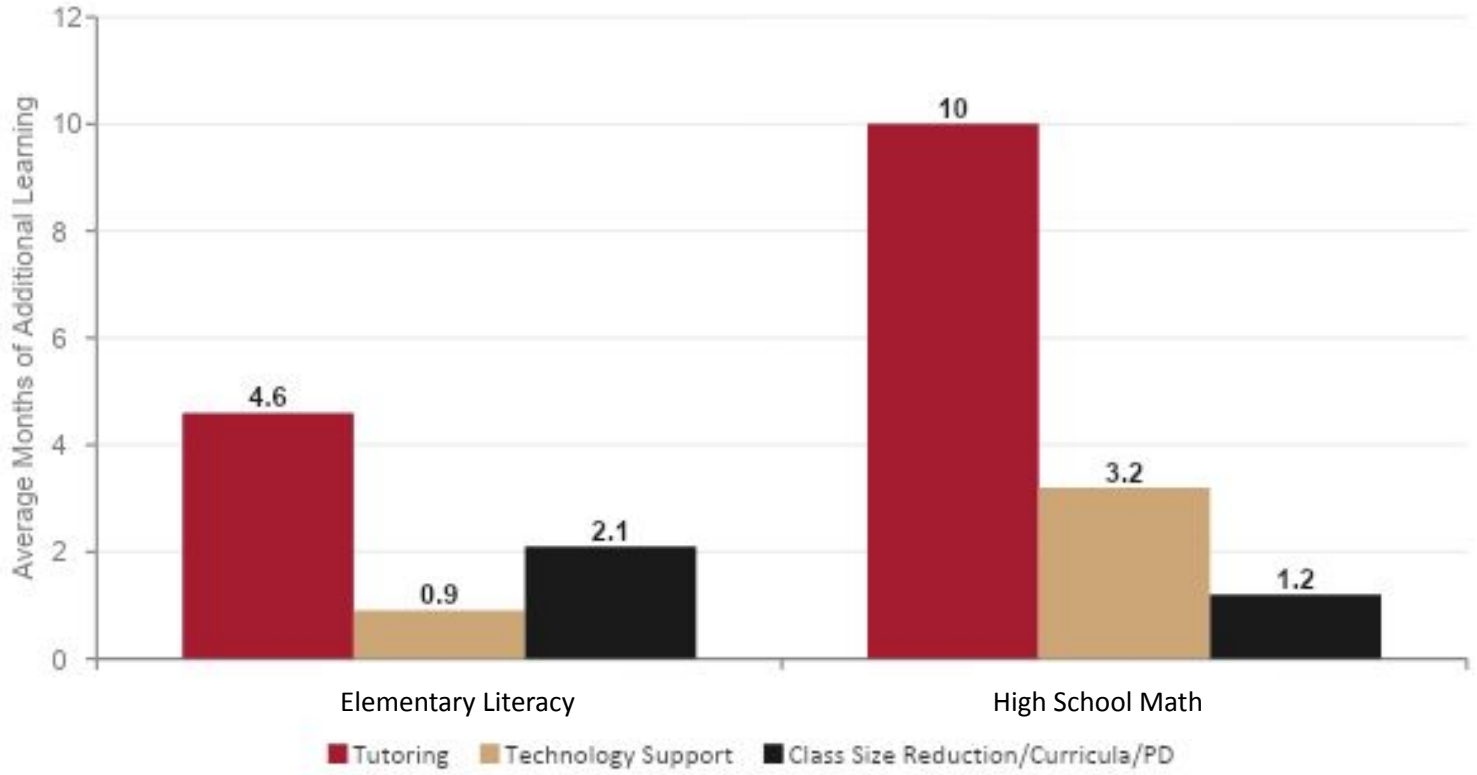
"If non-CTE graduates earn $40,000, CTE concentrators earn about $41,200 — $1,200 more per year, or $12,000 over a decade."

Reader: Okay, $1,200 a year — I can picture that.

MOVE: *Pick two realistic comparison points, report the predicted outcome at each, and let the reader see the gap themselves*

2 Compare effect sizes to other, more familiar or relatable effects

High-impact tutoring is more effective than other academic interventions.



Move:

Most readers can't tell you whether an effect size of 0.15 is impressive or trivial — and they shouldn't have to. Show them a comparison instead, so they can judge whether the number is big or small.

Robinson, Carly D. and Susanna Loeb. (2021). High-Impact Tutoring: State of the Research and Priorities for Future Learning (EdWorkingPaper: 21-384). Retrieved from Annenberg Institute at Brown University: https://edworkingpapers.com/ai21-384

A short checklist for your next draft

1. Open by stating the finding, the implication, and why the reader should care.
2. Move your most striking sentence to paragraph one.
3. Pick the calibrated middle: honest about what you found, accessible to the decider.
4. Don't write "mixed" when there is a helpful pattern.
5. Translate key numbers into something comparative, familiar, or human-scale.
6. Replace coefficients with predicted values or comparisons to other effect sizes

Thank you!

Christina Claiborne

christina_claiborne@brown.edu

Carrie Conaway

carrie_conaway@gse.harvard.edu

EdResearch for Action · Annenberg Institute at Brown University

Made to Stick	**Making Numbers Count**	**The WSJ Guide to Information Graphics**	**Better Presentations**
Chip & Dan Heath	*Chip Heath & Karla Starr*	*Dona Wong*	*Jonathan Schwabish*