

Improving the Usefulness & Use of Systematic Reviews of Evidence to Inform Policy & Practice

Conference on External Validity in Program Evaluation
Co-sponsored by the Laboratory for Interdisciplinary Evaluation of Public
Policies at Sciences Po & the University of Maryland School of Public
Policy
Paris, France
June 6, 2023

Rebecca A. Maynard
University Trustee Chair Professor Emeritus
University of Pennsylvania

Pre-conference Version 6-4-23

Outline

- ❑ Review evidence needs to support policy & practice
- ❑ Preview strengths of current infrastructure to support evidence informed policy-making
- ❑ Share examples of improvements in evidence pipeline & its use
- ❑ Share thoughts on room for improvement

Evidence Needs by User Group

Developers

- Product design & development
- Product marketing
- Efficient & adaptive implementation
- Product improvement

Consumers

- Policy making
- Product/practice selection
- Efficient & adaptive implementation

Diverse Methods for Meeting Evidence Needs

Developers

- Realist reviews
- Single study reviews
- Planned iterative study reviews (impact & implementation/process)
- Meta-Analysis of impacts across contexts
- Meta-regression focused on context variation
- Gap maps

Consumers

- Systematic reviews of evidence of impacts on outcomes, cost & cost-effectiveness
- Meta-Analysis of across contexts
- Gap maps

Strengths of Current Infrastructure

Developers

Statistical tools & training

- ☐ Common guidelines for education research & development
- ☐ Evidence review guidelines
- ☐ Guidelines on replication & reproducibility

Interdisciplinary sharing opportunities

- ☐ OSF, open access on-line training
- ☐ Growing culture of transparency & access
- ☐ Open access journals, study registration, data sharing hubs

Consumers

Evidence review platforms

- ☐ Single studies & syntheses (e.g., Campbell Collaboration; WWC, CLEAR, HomVEE) of qualitative & quantitative outcomes)
- ☐ BCA analyses (e.g., WSIPP; PowerUp!-CEA)

Routinized attention to needs

- ☐ Mandated gap maps & Evidence agendas for state & federal agencies
- ☐ Increased access to academic journals

Example 1: A Realist Review

Improving skills & care standards in the support workforce for older people (Rycroft-Malone et al., 2016)

Headline: “. . .using eight context–mechanism–outcome configurations, . . . the design and delivery of workforce development in older people’s care should consider [as] starting points, [supporting] workers’ personal factors, . . . requirements of workforce development [&} fit with broader organisational goals.

Aim: Synthesize evidence on whether a particular theory of change/logic model “fits” using informed stakeholders.

Product: “. . . an explanatory account of how the design and delivery of workforce development interventions work to improve the skills & care standards of support workers . . .”

Info on: Outcomes evaluated, client characteristics & services

Applications: (1) Innovative programs in use & (2) strategies for priority circumstances

Example 2: Typical Evidence Clearinghouse Review

Pathways to Work Evidence Clearinghouse (USDHHS)

Impact Evaluations: RCTs or QEDs

Aim: Synthesize evidence to on the effectiveness of workforce development programs that have been evaluated.

N = 305 studies of 231 interventions

Info on: Outcomes evaluated, client characteristics & services

Applications: Developers looking for ideas to build on; Consumers in search of strategies for improvement











Case Studies: (1) Innovative programs in use & (2) strategies for priority circumstances

Aim: Accelerate development; guide implementation and improvement work

Applications: Developers working to innovate or improve; Consumers looking for “fresh” ideas

[Link](#) to protocol for case study reviews

Example 2b: Pathways Review of Job Corps-Studies with Long-term Results

Outcome domain	Term	Effectiveness rating	Effect in 2018 dollars and percentages	standard deviations	Sample size
Increase earnings	Short-term		↓ -\$209 per year	-0.010	15138
	Long-term		↓* -\$84 per year	-0.004	15138
	Very long-term		↓ -\$42 per year	-0.002	15138
Increase employment	Short-term		↑ 4% (in percentage points)	0.102	15138
	Long-term		↑ 1% (in percentage points)	0.014	15138
	Very long-term		↓ -1% (in percentage points)	-0.019	15138
Decrease benefit receipt	Short-term		↑ -\$85 per year	-0.031	11641
	Long-term		↑ -\$47 per year	-0.017	11313
	Very long-term		↑ -\$58 per year	-0.021	15127
Increase education and training	All measurement periods		↑ 9% (in percentage points)	0.179	11313

Example 2b: Pathways Clearinghouse- Case Studies

Purpose: A companion to the Pathways Clearinghouse which has reviewed more than 300 studies

Information provided:

Intervention description

Outcomes evaluated

Client characteristics & services

Corpus: 10 Case Studies: (1) Innovative programs in use & (2) strategies for priority circumstances

Aim: Provide practitioners with information about the experiences of other organizations implementing innovative interventions.

[Link](#) to protocol

Example 3: Meta-Regression

Meta-regression of study findings on programs for English Language Learners (Williams & Garrett 2023)

Purpose: “Learn what works, where and for whom.”

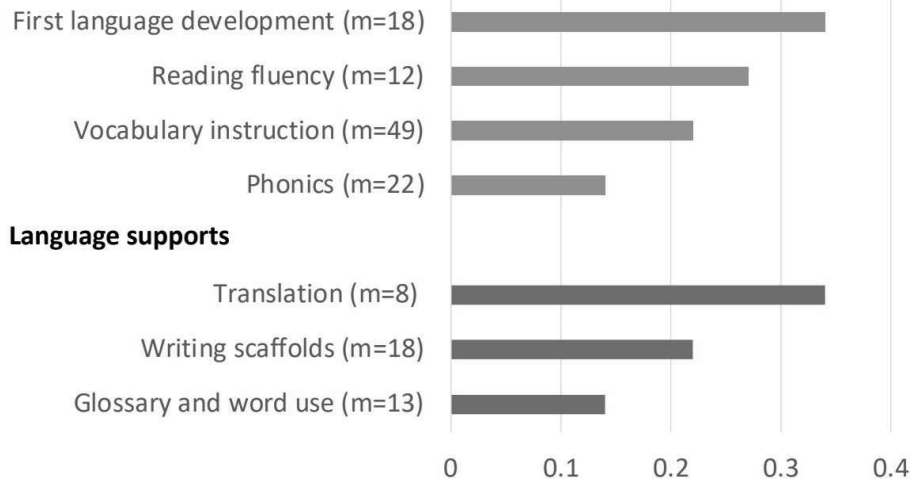
Study Questions: Questions: How are program impacts related to outcomes, program features, school characteristics, student characteristics & study measures and design?

Sample: N = 79 of 826 full-text studies reviewed

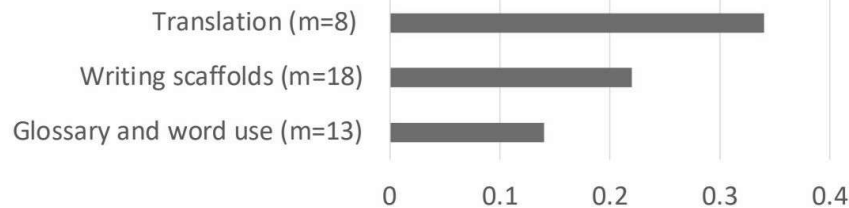
Example 3a: Meta-Regression- Sample Output

Impacts by approach and language supports

Approaches to language development

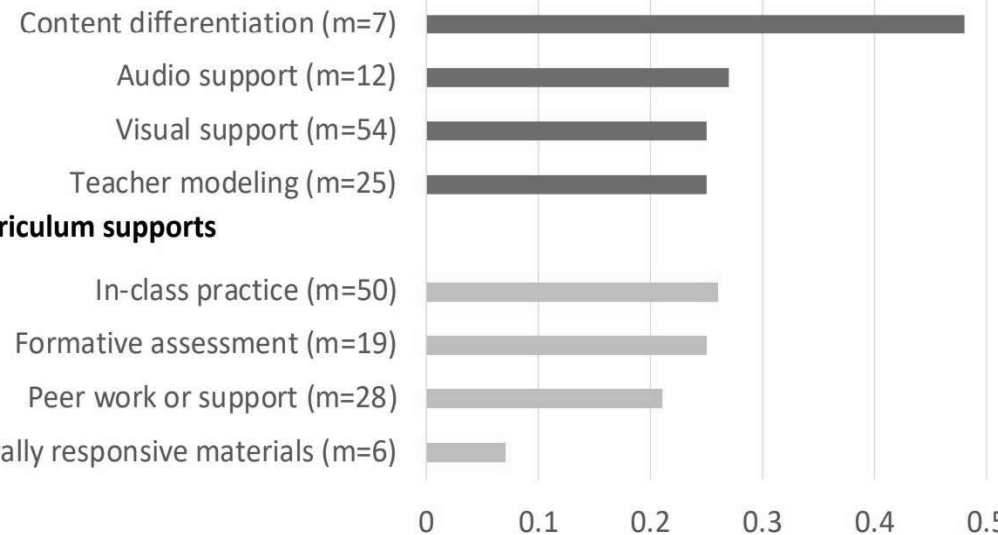


Language supports

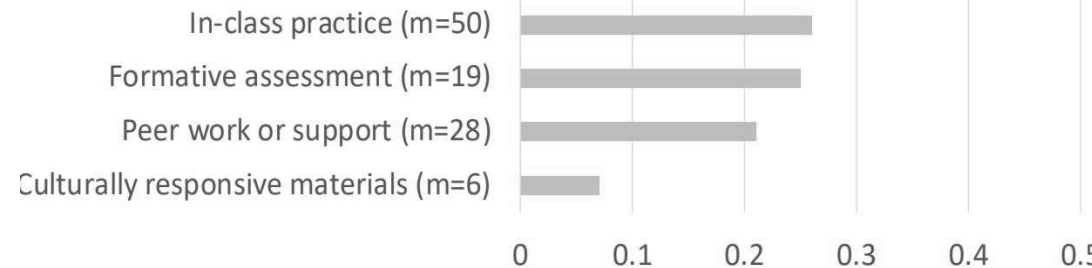


Weighted Average Effect Size, in Standard Deviations

Content accessibility supports



Curriculum supports



Example 4: THE-RCT- A Repository of Harmonized Micro Data

Synthesis of micro-data from prior studies of a class of interventions (Weiss & Bloom 2022)

Purpose: Use cross-site variation in intervention design & targeting to improve understanding of what work, for whom, & under what conditions

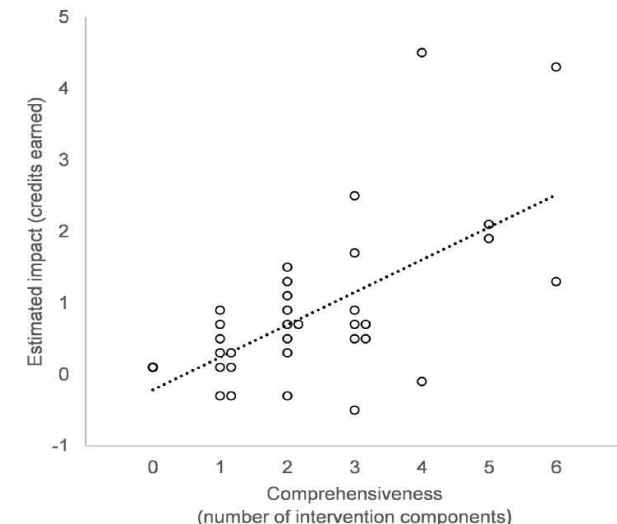
Study sample: RCTs of 39 post secondary interventions and a sample 60,000 individuals

Findings: Micro-regression analysis of impacts and impact variation associated with program characteristics

PREVALENCE AND NUMBER OF COMPONENTS
ACROSS INTERVENTIONS

INTERVENTION COMPONENTS	PERCENTAGE OF INTERVENTIONS
Presence of component	
Financial support	51
Enhanced advising	38
Tutoring	28
Learning communities	23
Success course	23
Promotion of full-time/summer enrollment	33
Instructional reform	26
Comprehensiveness (number of components)	
0	3
1	38
2	23
3	21
4	5
5	5
6	5
Number of interventions	39

INTERVENTION IMPACTS ON CREDITS EARNED VERSUS
THEIR COMPREHENSIVENESS (YEAR 1)



Rooms for Improvement

- 1. Design evidence-building agendas in anticipation of pooling findings across studies**
Develop, use & report the prevailing theory of change
Where possible, harmonize measurement
- 2. Align study design and measurement plans with study context & theory of change**
Be strategic in defining populations, interventions, comparisons, outcomes & timeframes (PICOTs) for the study
Include intentionally designed implementation research (e.g., building on theory of change/logic model)
Plan to measure costs & benefits for various stakeholder groups
- 3. Pre-register studies & update registrations as warranted**
Provide open access reporting of study design, characteristics & findings
Plan for respectful sharing of micro-data
- 4. Fully report study design, analysis methods & findings outside of paywalls**
- 5. Harmonize tools methods and tools for research synthesis & methods across disciplines**
- 6. Improve the usability & use of evidence repositories**

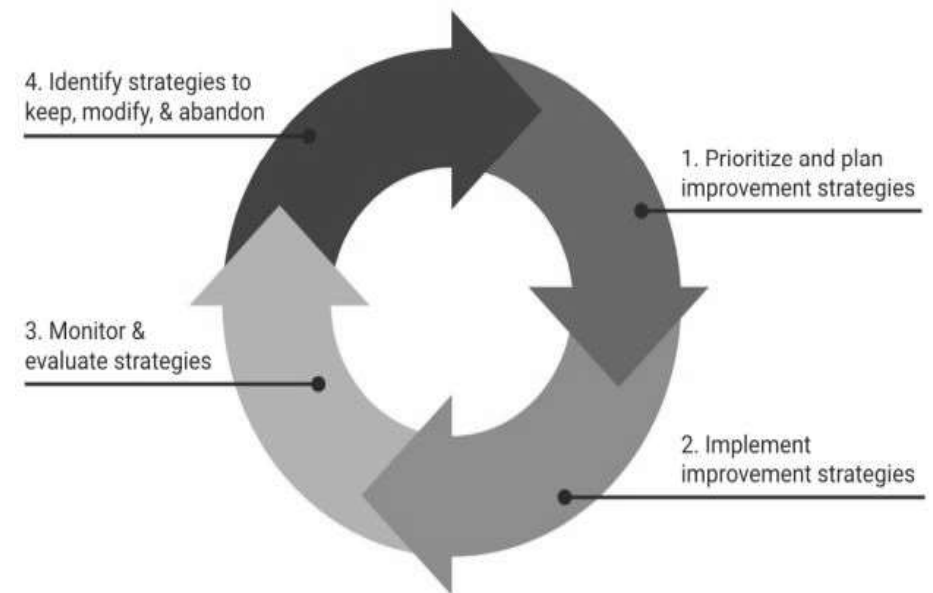
Anchor Qualities of Useful & Used Evidence

3-way Engagement: Research, Developer, Consumer



Ongoing Feedback Loops

Example: Year Up's Workforce Development Initiative



Sampling of Resources

Study design & analysis

Common guidelines for education research & development

Companion guidelines on replication & reproducibility [Link](#)

Research synthesis methods-

Campbell Collaboration [Link](#)

What Works Clearinghouse Handbook [Link](#)

Generalizability: The Generalizer [Link](#)

Cost-benefit analysis: CBCSE [Link](#); EPA [Link](#); Social Programs [Link](#)

Power analysis [Link](#)

Research synthesis methods

MetaReviewer [Link](#)

Evidence gap map builder [Link](#)

Evidence review repositories

WSIPP [Link](#)

What works clearinghouse [Link](#)

Clearinghouse for Labor Research [Link](#)

Home Visiting Evidence of Effectiveness [Link](#)

Partial Resource Wish List

Citation archive/registry

Registry of reviewed studies

Citation & meta data for all studies reviewed for systematic review or evidence platforms

Evidence gap maps

Build-a-Gap-Map tool

Uses data in the citation archive

Computational tools

“Common MA” software

Flexible data input format **but uniform data archive format**

Data archives

Shared data archive in uniform format for reviewed studies

Doi & credit for original coders

References & Resources

- Britt, J., Fein, D., Maynard, R. & Warfield (2021). Improving Academic Success and Retention of Participants in Year Up's Professional Training Corps. Project Evident. <https://projectevident.org/wp-content/uploads/2022/09/YearUpActionableEvidenceCaseStudyJuly21.pdf>
- Cost Analysis Standards Project. (2021). Standards for the economic evaluation of educational and social programs. American Institutes for Research. <https://www.air.org/sites/default/files/Standards-for-the-Economic-Evaluation-of-Educational-and-Social-Programs-CASP-May-2021.pdf>
- Cramer, J. (2022). WSIPP's LAP inventory: A brief history and potential changes for the future (Document Number 22-05-2201). Olympia: Washington State Institute for Public Policy https://www.wsipp.wa.gov/ReportFile/1746/WSipp_WSIPPs-LAP-Inventory-A-Brief-History-and-Potential-Changes-for-the-Future_Report.pdf
- Gardiner, K. and R. Juras. (2019). PACE Cross-Program Implementation and Impact Study Findings, OPRE Report #2019-32, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. https://www.acf.hhs.gov/sites/default/files/documents/opre/pace_cross_program_implementation_and_impact_study_findings_final.pdf
- Peck, Laura R., Schwartz, Deena, Strawn, Julie, Weiss, Christopher C., Juras, Randall, Mills de la Rosa, Siobhan, Greenstein, Nathan, Morris, Tori, Durham, Gabriel, & Lloyd, Charlotte. (2021). A Meta-Analysis of 46 Career Pathways Impact Evaluations. Report prepared for the U.S. Department of Labor, Chief Evaluation Office, Rockville, MD: Abt Associates. https://www.abtassociates.com/files/insights/reports/2022/a-Meta-Analysis-of-46-career-pathways-impact-evaluations_final-report.pdf
- Rycroft-Malone, J., Burton, C., Hall, B., McCormack, B., Nutley, S., Seddon, D., & Williams, L. (2014). Improving skills and care standards in the support workforce for older people: a realist review. *BMJ open*, 4(5), e005356. <https://bmjopen.bmj.com/content/bmjopen/4/5/e005356.full.pdf>
- Stone et al. What works to improve early grade literacy in Latin America and the Caribbean? A systematic review and Meta-Analysis (2020).
- Campbell Systematic Reviews, Vol. 16, Issue 1. <https://onlinelibrary.wiley.com/doi/epdf/10.1002/cl2.1067>
- Strawn, Julie, Laura R. Peck & Deena Schwartz (2021). New Insights on Career Pathways: Evidence from a Metanalysis. Washington, DC: U.S. Department of Labor, Chief Evaluation Office. https://www.abtassociates.com/files/insights/reports/2022/new-insights-on-career-pathways_evidence-from-a-Meta-Analysis_summary-brief.pdf
- Tipton, E. (2014) How generalizable is your evaluation? Comparing a sample and population through a generalizability index. *Journal of Educational and Behavioral Statistics*, 39(6): 478 – 501.
- Tipton, E. & Miller, K. (2023) The Generalizer. Webtool hosted at <https://thegeneralizer.org>.
- Weiss, M. J., Bloom, H. S., & Singh, K. (2022). What 20 Years of MDRC RCTs Suggest About Predictive Relationships Between Intervention Features and Intervention Impacts for Community College Students. *Educational Evaluation and Policy Analysis*, 01623737221139493. <https://journals.sagepub.com/doi/pdf/10.3102/01623737221139493>
- Williams, R. & R. Garrett. Understanding Variation in Program Impacts For English Learners: What Does the Research Say? Preliminary Findings and Implications for Practice. Presented at the National Association of Bilingual Education (NABE) Annual Conference | January 2023. <https://www.air.org/sites/default/files/2023-02/Understanding-Variation-in-Program-Impacts-for-ELs-NABE-Presentation-January-2023.pdf>
- Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>