Equity-Related Moderator Analysis in Syntheses of Dependent Effect Sizes: Conceptual and Statistical Considerations

James E. Pustejovsky, Jingru Zhang, & Beth Tipton

Society for Research on Educational Effectiveness
September 27, 2023
Arlington, VA

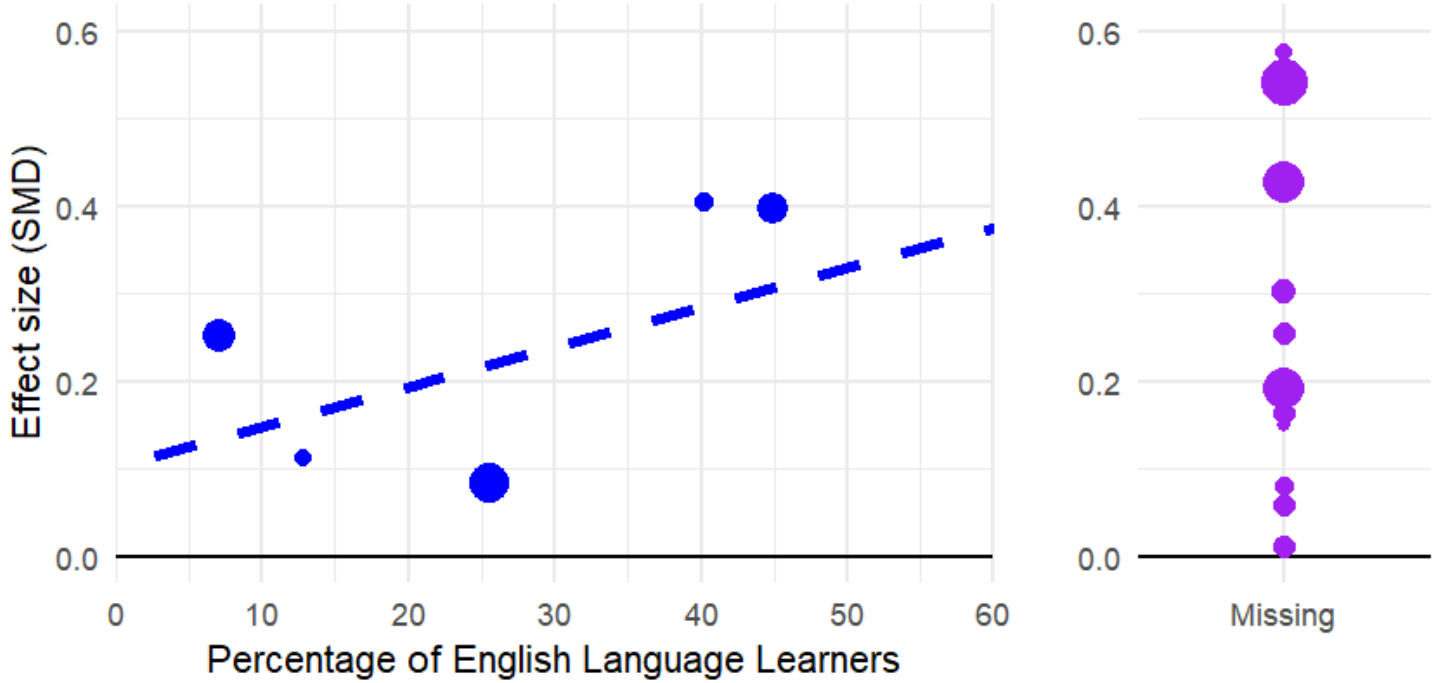
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Equity-related moderator analysis

- In syntheses of educational intervention studies, our goal is to understand the **distribution of program impacts**.
    - Equity-related moderator analyses seek to address questions of **who benefits** from an intervention and **how benefits and harms are distributed** across students.
- Moderator analyses examine variation in effect size based on characteristics of primary study participants and contexts:
    - Participants' family income level
    - Participant racial/ethnic groups
    - Participant English Language Learner status
    - School urbanicity

Synthesis of study-level average effects

- Traditional synthesis involves examining associations between average effect sizes and aggregate sample characteristics.



Source: Neitzel, et al. (2020). Data archive for "Success for All: A Quantitative Synthesis of U. S. Evaluations." Towson, MD: Center for Research and Reform in Education (CRRE), Johns Hopkins University. Retrieved from https://github.com/aj-neitzel/Success-for-All-A-Quantitative-Synthesis-of-U.-S.-Evaluations

Synthesis of dependent effect sizes

- Contemporary syntheses often involve **multiple effect size estimates** from some or all studies.

- Results from each of *multiple samples*
- Results on *multiple outcome measures*
- Results at *multiple follow-up times*
- Results for each of *several subgroups*

Study	Sample	Subgroup	Followup	ELL %	N	ES 1	ES 2	ES 3
A	A.1	Non-ELL	Short	0	108	0.27	0.66	0.32
A	A.1	ELL	Short	100	21	0.74	0.14	0.47
B	B.1	Non-ELL	Short	0	48	-0.42	0.24	0.23
B	B.1	ELL	Short	100	36	-0.24	-0.16	0.18
B	B.1	Non-ELL	Long	0	48	0.63	0.45	0.02
B	B.1	ELL	Long	100	36	-0.37	0.49	0.30
C	C.1	Mix	Short	20	77	0.32	0.24	0.02
C	C.2	Mix	Short	20	46	0.44	-0.12	0.62
C	C.3	Mix	Short	25	52	-0.42	0.07	0.50
D	D.1	Mix	Short	22	114	0.14	0.27	0.40
D	D.1	Mix	Long	22	114	0.55	0.19	0.26
D	D.2	Mix	Short	35	97	0.10	0.42	0.22
D	D.2	Mix	Long	35	97	-0.01	0.07	-0.15

Direct evidence

- Reported effect size estimates for each of multiple subgroups.
- Provides estimates of ***individual-level variation*** in impacts.
- Study-level operational features are held constant.

Study	Followup	ELL %	N	ES 1	ES 2	ES 3
A	Short	0	108	0.27	0.66	0.32
A	Short	100	21	0.74	0.14	0.47
B	Short	0	48	-0.42	0.24	0.23
B	Short	100	36	-0.24	-0.16	0.18
B	Long	0	48	0.63	0.45	0.02
B	Long	100	36	-0.37	0.49	0.30

Contextual evidence

- **Sample-level average** effect size estimates and **average sample characteristics**.
- Open to **aggregation bias** (a.k.a. the ecological fallacy).

Study	Sample	Followup	N	ELL %	ES 1	ES 2	ES 3
A	A.1	Short	129	16.28	0.35	0.58	0
B	B.1	Long	84	42.86	0.20	0.46	0
B	B.1	Short	84	42.86	-0.34	0.07	0
C	C.1	Short	77	20.00	0.32	0.24	0
C	C.2	Short	46	20.00	0.44	-0.12	1
C	C.3	Short	52	25.00	-0.42	0.07	0
D	D.1	Long	114	22.00	0.55	0.19	0
D	D.1	Short	114	22.00	0.14	0.27	0
D	D.2	Long	97	35.00	-0.01	0.07	0
D	D.2	Short	97	35.00	0.10	0.42	0

Direct and contextual evidence are *conceptually distinct...*

...and should be analyzed as such.

- Meta-analyze the direct evidence (subgroup-specific effect sizes) alone, excluding the contextual evidence.

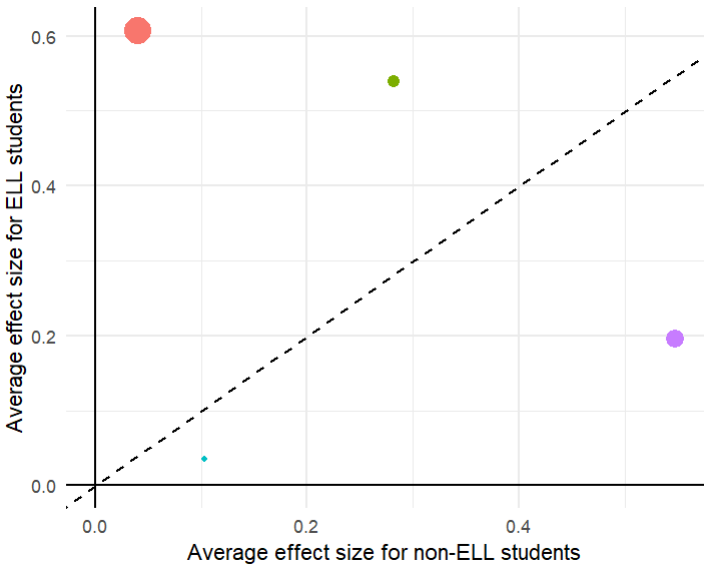
and/or

- Center the predictor by sample, include the centered predictor and the sample-level averaged predictor in a meta-regression.

Meta-analyze the direct evidence alone

- Analyze the direct evidence (subgroup-specific effect sizes) in a separate meta-analysis, excluding the contextual evidence.

$$\begin{pmatrix} ES_j^{non} \\ ES_j^{ELL} \end{pmatrix} = \begin{pmatrix} \mu_{non} \\ \mu_{ELL} \end{pmatrix} + \begin{pmatrix} v_{0j} \\ v_{1j} \end{pmatrix} + \begin{pmatrix} e_{0j} \\ e_{1j} \end{pmatrix}$$



Center by sample

- Calculate sample-level aggregate characteristic for each unique sample:

$$\left(\overline{ELL\%}\right)_j = \frac{1}{\sum_{i=1}^{k_j} N_{ij}} \sum_{i=1}^{k_j} N_{ij} \times (ELL\%)_{ij}$$

- Estimate a meta-regression with sample-centered and sample-aggregate predictors:

$$ES_{ij} = \beta_0 + \beta_1 \underbrace{\left(ELL\%_{ij} - \overline{ELL\%}_j\right)}_{\text{direct evidence}} + \beta_2 \underbrace{\left(\overline{ELL\%}\right)_j}_{\substack{\text{contextual} \\ \text{evidence}}} + v_{ij} + e_{ij}$$

  - $\hat{\beta}_1$ is based only on samples providing direct evidence
  - $\hat{\beta}_2$ is based on sample-level aggregated effect sizes

Current practice

- We reviewed empirical meta-analysis projects funded by the Institute of Education Sciences between 2002 and 2018.
- 25 projects included "meta-analysis" in project description and had associated journal article reporting a meta-analysis.

Feature	Category	N	Pct
Any moderator analysis		24	96
Student characteristic moderators		16	64
Centering	Grand-mean	3	12
	Sample-mean	1	4
	Not specified	1	4
Working model	Correlated effects	9	36
	Aggregated effects	7	28
	Hierarchical effects	3	12
	Independent effects	2	8
	Multi-level	2	8

Further Recommendations

- Prior to conducting moderator analysis, **describe the structure of the evidence** on equity-related student characteristics.

Variable	Reported N ES (%)	Reported N Studies (%)	M	SD	Within-Study Variation N Studies (%)
Grade	1061 (0.96)	176 (92)	3.32	2.93	26 (14)
Male Pct	777 (0.70)	124 (65)	0.52	0.14	45 (32)
White Pct	656 (0.59)	109 (57)	0.40	0.27	41 (31)
Economic Disadvantage Pct	462 (0.42)	77 (40)	0.57	0.24	27 (28)
ELL Pct	385 (0.35)	56 (29)	0.22	0.24	23 (35)
SPED Pct	316 (0.28)	48 (25)	0.20	0.28	19 (33)

- If student characteristics are of focal interest, **use data extraction strategies to maximize amount of direct evidence.**

Limitations and future directions

- Data availability is a major limitation
  - Common to have missing information about sample-average characteristics.
  - Subgroup-specific results available only for a small subset of studies.
- Selective reporting of subgroup analysis could create biases in direct evidence (Hahn et al., 2000).
- Need to further develop working models for synthesizing direct and contextual evidence together.