2-17-2017

Designing Impact Studies in Education in the Past 15 Years: Have We Made Any Progress?

Jessaca Spybrook
Western Michigan University, jessaca.spybrook@wmich.edu

Designing Impact Studies in Education in the Past 15 Years: Have We Made Any Progress?

Jessaca Spybrook

February 17, 2017

*Joint work with Ran Shi and Ben Kelcey

Outline

- Brief History
- Design of Impact Studies
- Progress in Design?
- Next Steps

Brief History

Brief History

- 15 years ago
  - Institute of Education Sciences (IES)
    - Research branch of US Department of Education
  - Support rigorous research, evaluation, and statistics to improve education
  - Four centers to address these different goals

Brief History

- National Center for Education Research (NCER)
- Rigorous research
  - Foundational and Exploratory Research
  - Design and Development Research
  - Impact Research

Brief History

- National Center for Education Research (NCER)
- Rigorous research
  - Foundational and Exploratory Research
  - Design and Development Research
  - Impact Research

Brief History

- Impact Research
  - Seeks to establish rigorous evidence of effectiveness of an intervention
  - Answer “what works” question
  - Examples
  - In past 15 years, NCER funded more than 160 impact studies
  - Individual study > 3 million dollars

Design of Impact Studies

Design of Impact Studies

- Strong quasi-experiments
- Randomized trials prioritized
- Cluster randomized trials
  - Treatment assigned to entire clusters, or intact groups of individuals
  - Schools are common clusters
  - Outcomes measured at individual level

Design of Impact Studies

- Why cluster randomized trials?
  - Interventions often implemented at school level
  - Nested structure of schooling
  - Increase participation
  - Reduce contamination

Design of Impact Studies

- Common belief:
  - Presence of cluster randomized trial = rigorous evidence

Design of Impact Studies

- Common belief:
  - Presence of cluster randomized trial = rigorous evidence
- NOT TRUE

Design of Impact Studies

- Common belief:
  - Presence of cluster randomized trial = rigorous evidence
- NOT TRUE
- Important aspects of the:
  - Design
  - Implementation
  - Analysis

Design of Impact Studies

- Common belief:
  - Presence of cluster randomized trial = rigorous evidence
- NOT TRUE
- Important aspects of the:
  - Design
  - Implementation
  - Analysis

Design of Impact Studies

- Two important dimensions of the design of cluster randomized trial

- Dimension 1: Size of the study
  - Total number of clusters
  - Number of individuals per cluster

Design of Impact Studies

- Two important dimensions of the design of cluster randomized trial

- Dimension 1: Size of the study
  - Total number of clusters
  - Number of individuals per cluster

Design of Impact Studies

- Dimension 2: Precision of the study
  - Minimum detectable effect size (MDES)
  - Smallest true mean program effect size a study can detect for a given level of power

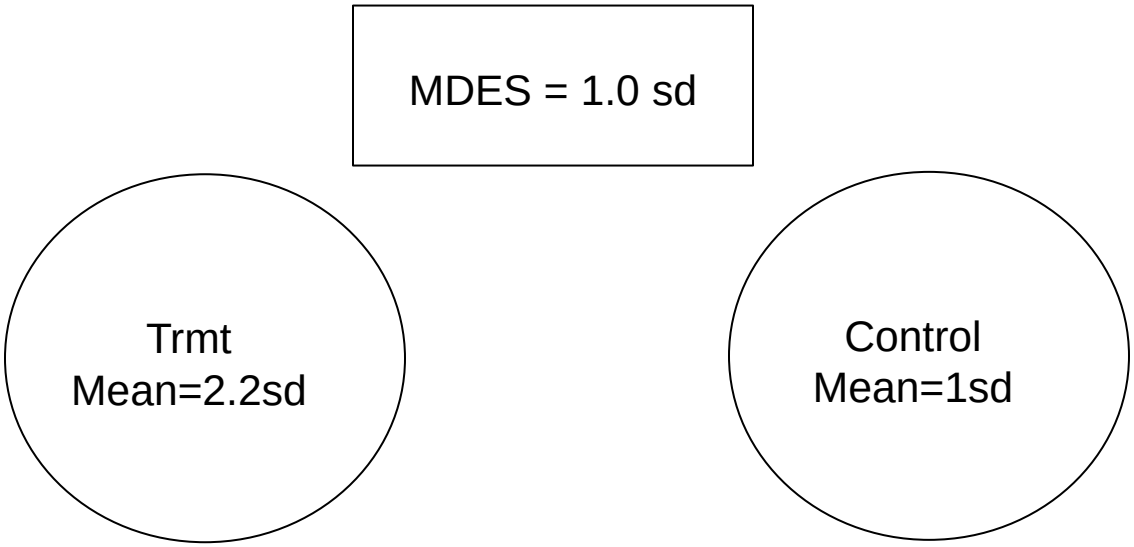
Design of Impact Studies

- Dimension 2: Precision of the study
  - Minimum detectable effect size (MDES)
  - Smallest true mean program effect size a study can detect for a given level of power

MDES = 1.0 sd

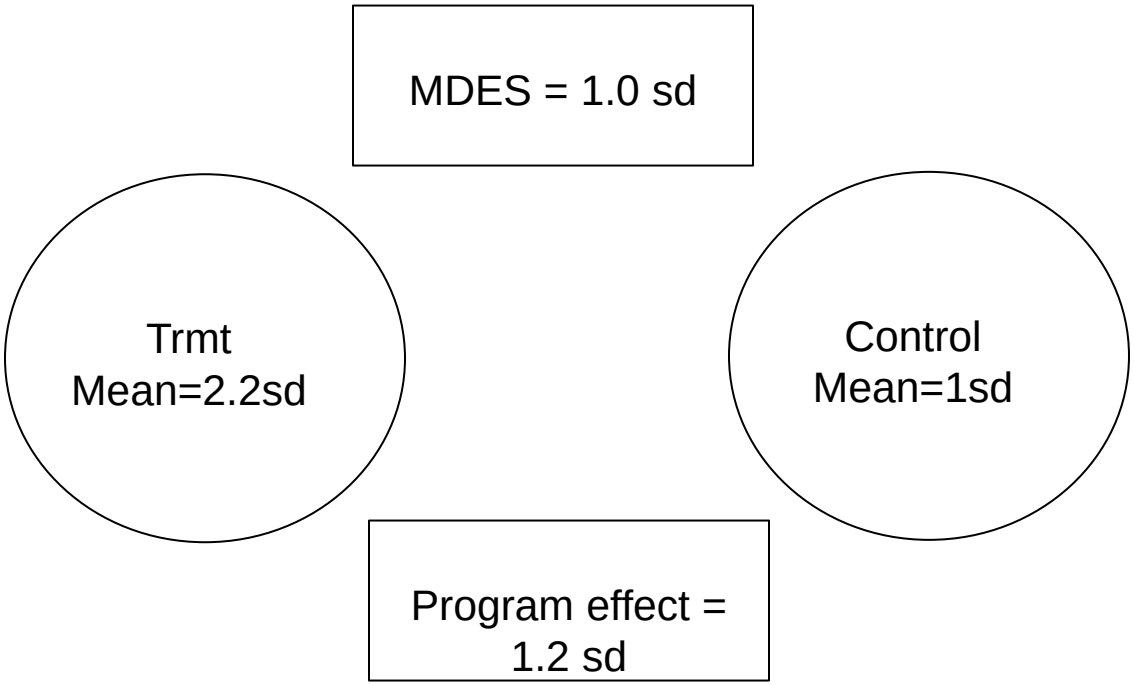
Design of Impact Studies

- Dimension 2: Precision of the study
  - Minimum detectable effect size (MDES)
  - Smallest true mean program effect size a study can detect for a given level of power



Design of Impact Studies

- Dimension 2: Precision of the study
  - Minimum detectable effect size (MDES)
  - Smallest true mean program effect size a study can detect for a given level of power



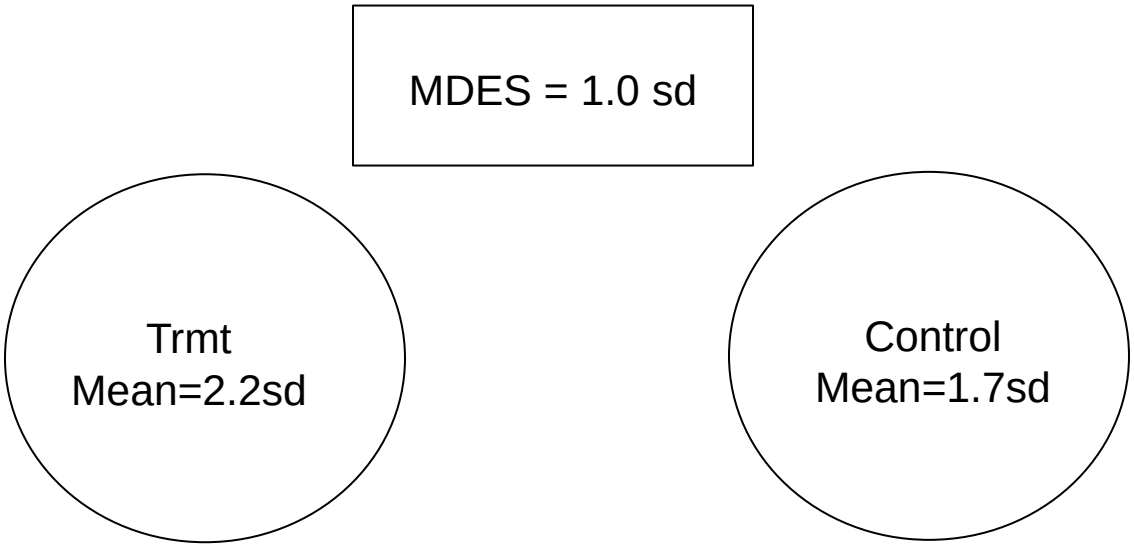
Design of Impact Studies

- Dimension 2: Precision of the study
  - Minimum detectable effect size (MDES)
  - Smallest true mean program effect size a study can detect for a given level of power

MDES = 1.0 sd

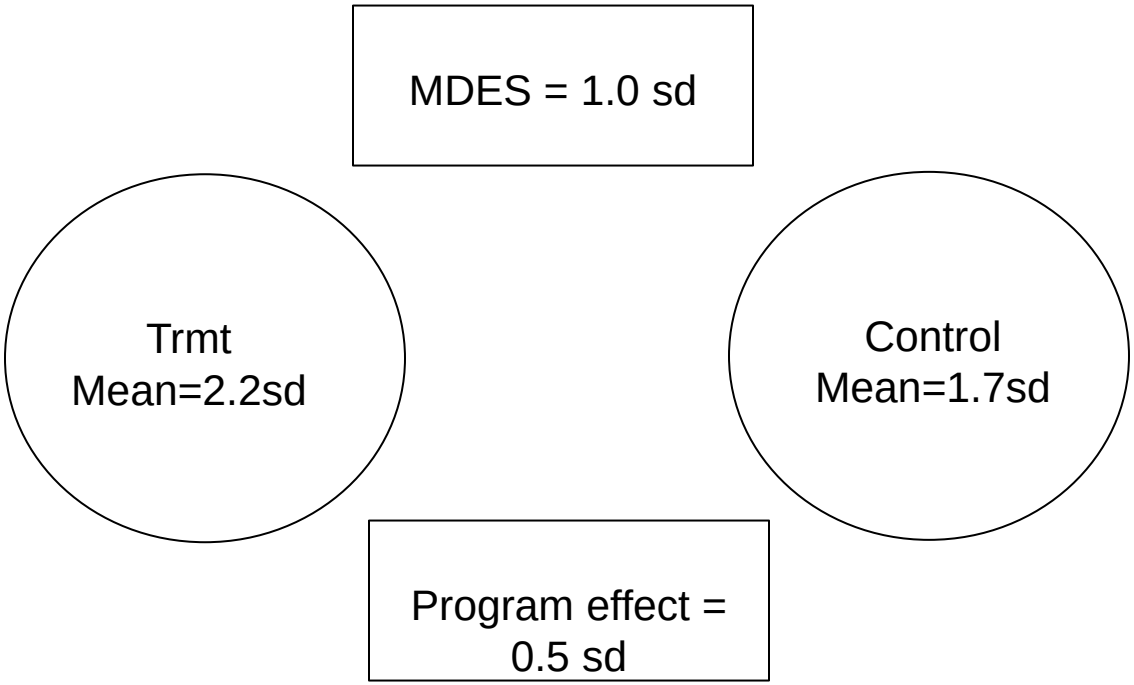
Design of Impact Studies

- Dimension 2: Precision of the study
  - Minimum detectable effect size (MDES)
  - Smallest true mean program effect size a study can detect for a given level of power



Design of Impact Studies

- Dimension 2: Precision of the study
  - Minimum detectable effect size (MDES)
  - Smallest true mean program effect size a study can detect for a given level of power



Design of Impact Studies

- Critical to design a study with a reasonable MDES

Design of Impact Studies

- Critical to design a study with a reasonable MDES

- What size treatment effects do we expect for educational interventions?
  - Context specific – depends on intervention, target grade, outcome type
  - Academic outcomes, meaningful range 0.20-0.30
    - Meta-analyses of intervention studies
    - Empirical estimates of average growth per year

Progress in Design?

Progress in Design?

- Compare designs of “early” studies to “recent” studies along the 2 design dimensions
  - Size of the study
  - Precision of the study

Progress in Design?

- Inclusion criteria
  - Impact trials, funded by NCER, use a cluster randomized trial, examine academic achievement
- Early Studies
  - Funded 2002-04
  - 16 studies
- Recent Studies
  - Funded 2011-2013
  - 22 studies

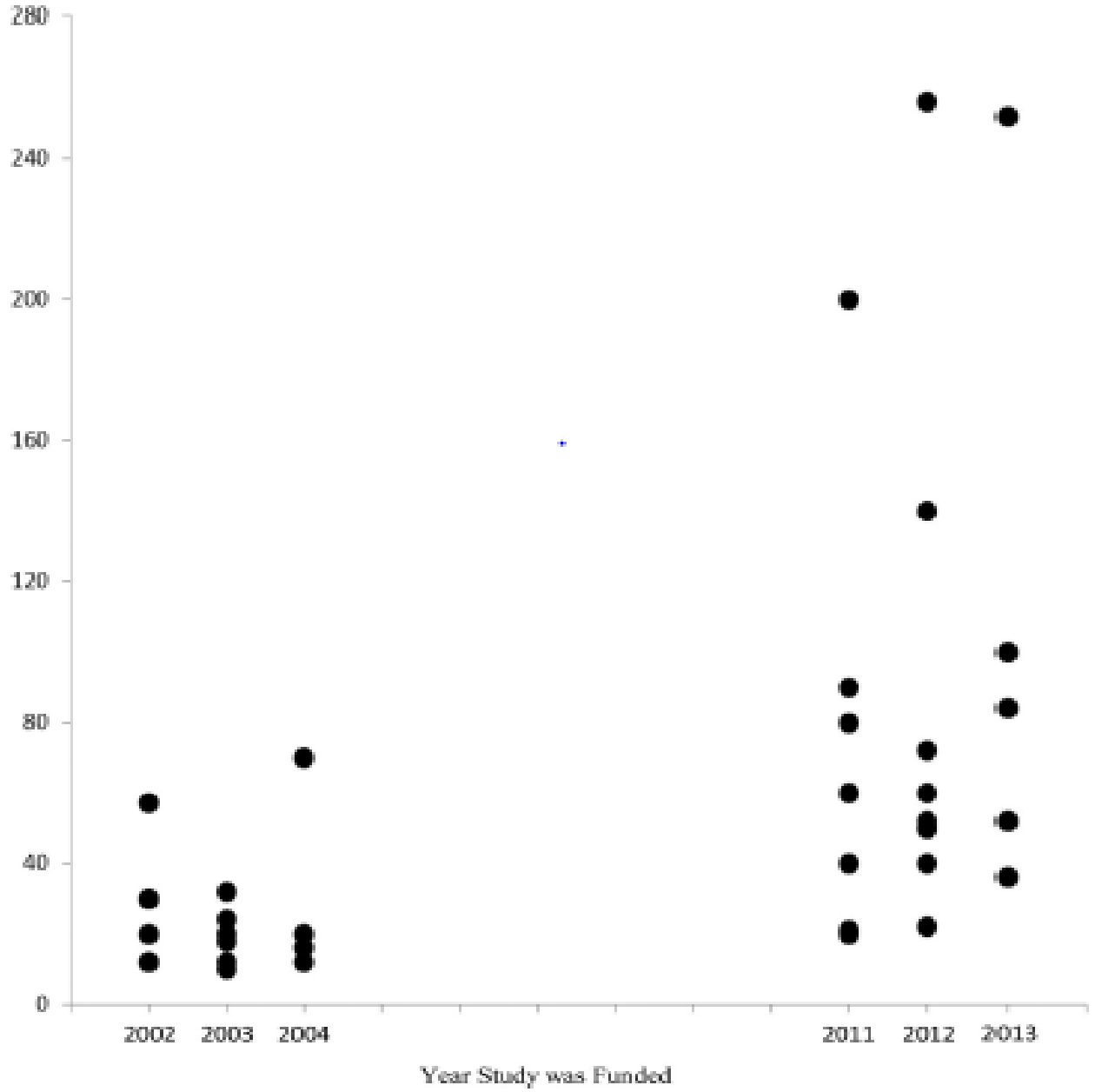
Progress in Design?

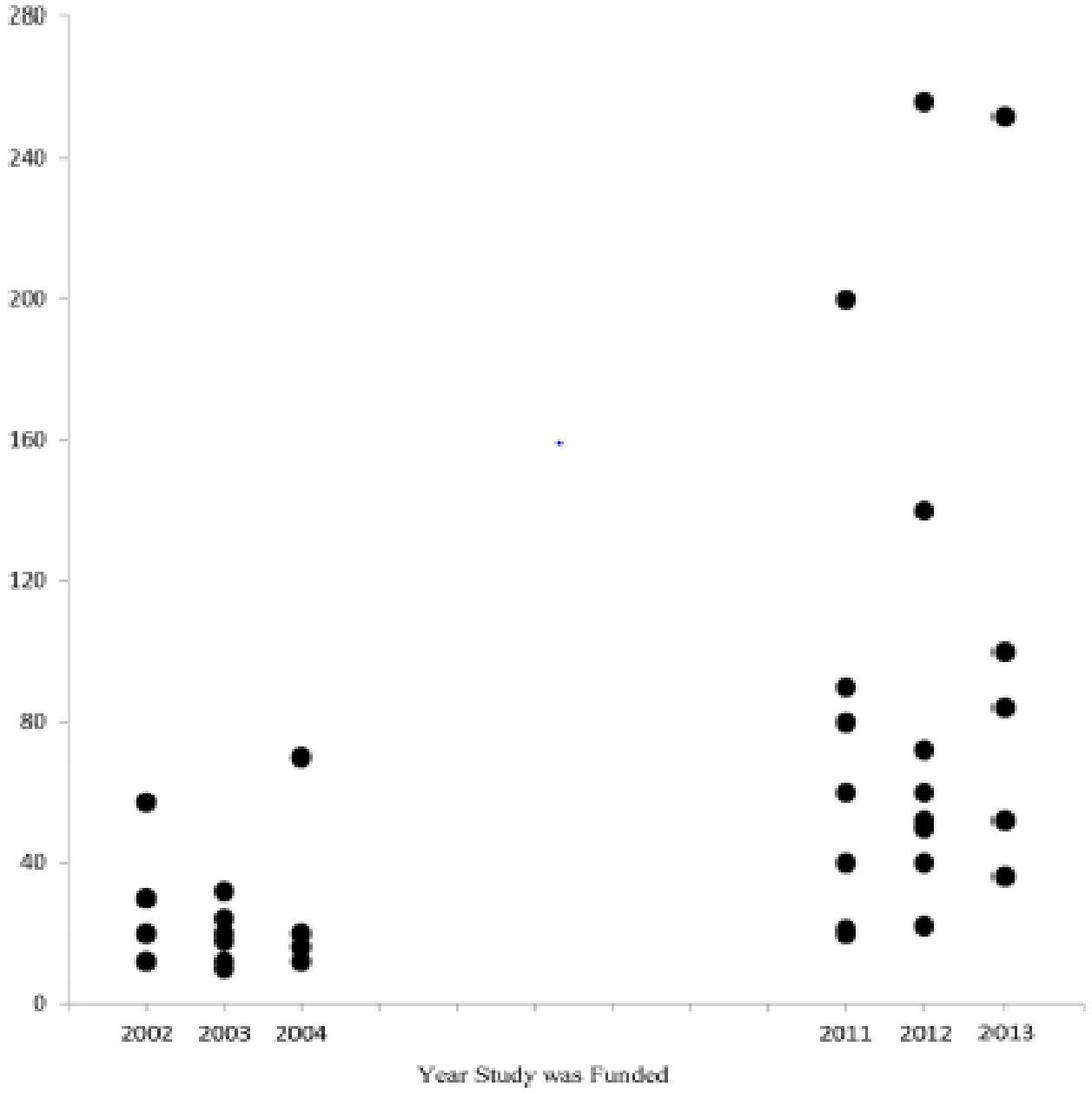
- Assess two dimensions of the design:
  - Size of study
    - Total number of clusters in the study
  - Precision of study
    - Minimum detectable effect size (MDES)

Progress in Design?

Dimension 1

Total Number of Clusters


280
240
200
160
120
80
40
0
2002
2003
2004
2011
2012
2013
Year Study was Funded



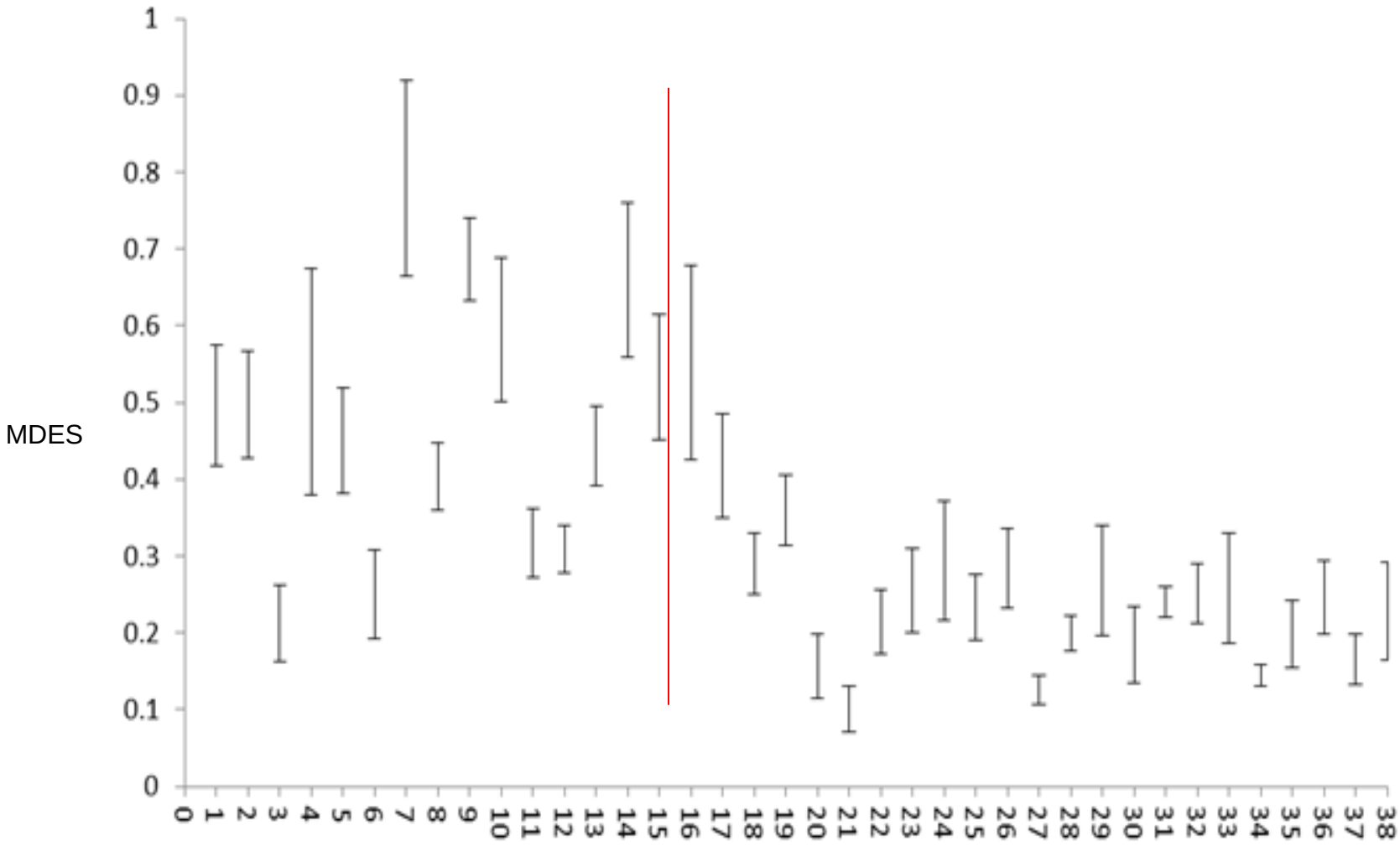
Early Studies
Median = 20

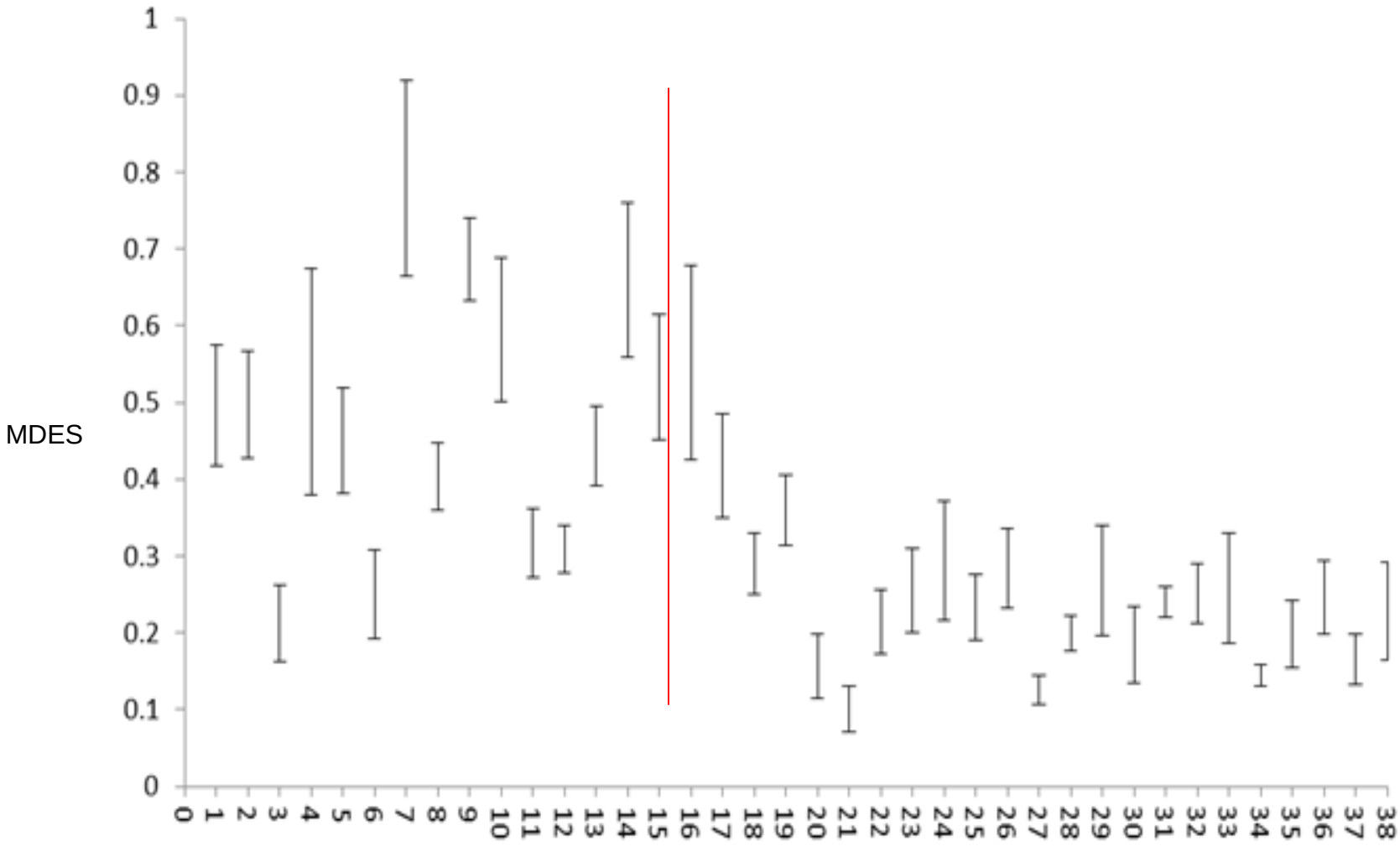
Recent Studies
Median = 52

Progress in Design?

Dimension 2

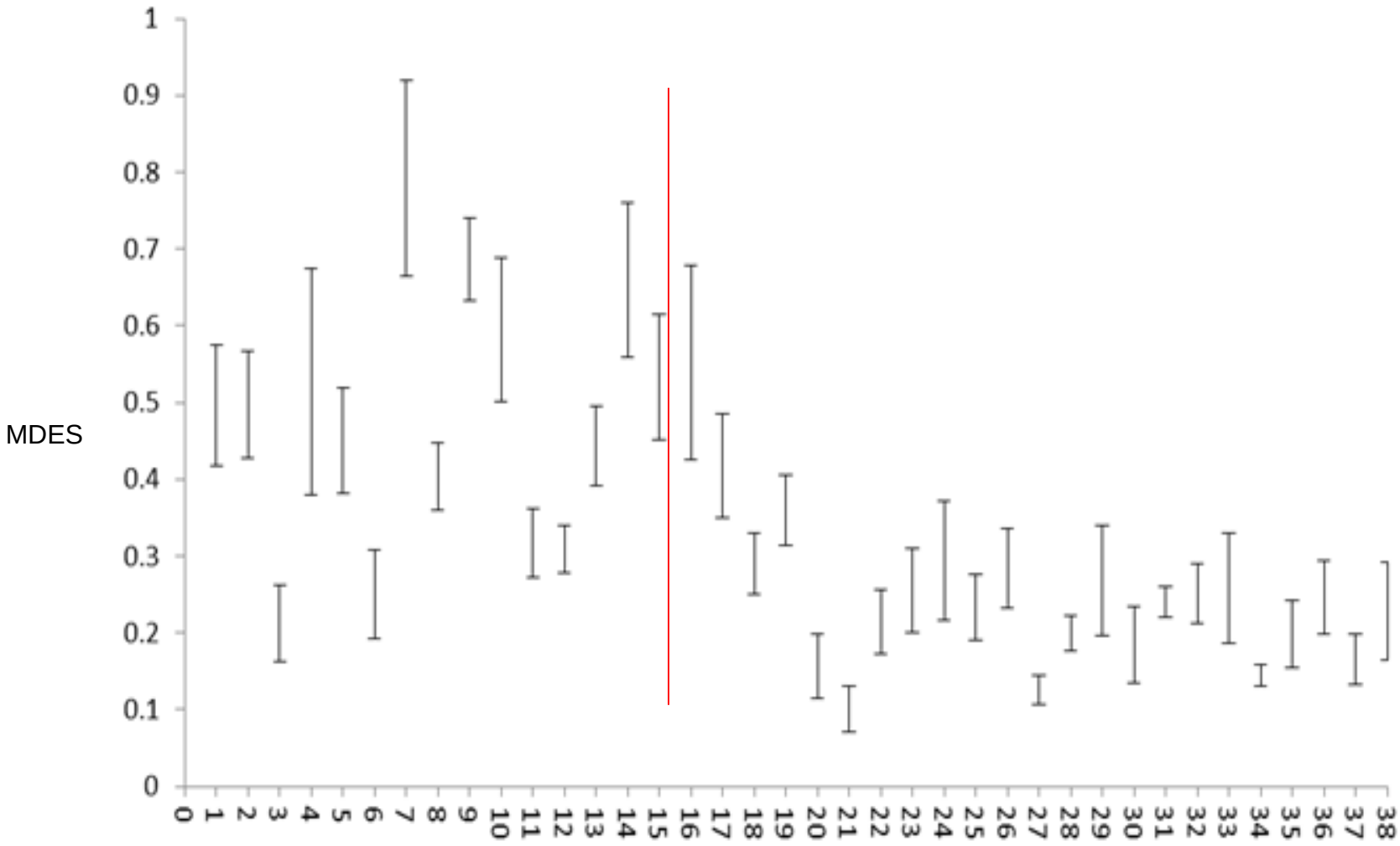
MDES


MDES
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38



Early Studies
Mean = 0.48

Recent Studies
Mean = 0.23



Early Studies
Mean = 0.48

t=5.81,p<.001

Recent Studies
Mean = 0.23

Progress in Design?

Progress in Design?

Yes!

Along these 2 design dimensions

Next Steps

Next Steps

- Expand the scope of questions
  - Design studies to detect *for whom* a program works
  - Design studies to detect *under what conditions* a program works

Next Steps

- Expand methodological work
  - Precision to detect moderator effects
  - User-friendly software to accompany tools

Questions?

jessaca.spybrook@wmich.edu

References

Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review*, 19(5), 547-556.

Spybrook, J., Shi,R., & Kelcey, B. (2016). Progress in the past decade: An examination of the precision of cluster randomized trials funded by the U.S. Institute of Education Sciences. *International Journal of Research and Method in Education, 39(3), 255-267*.