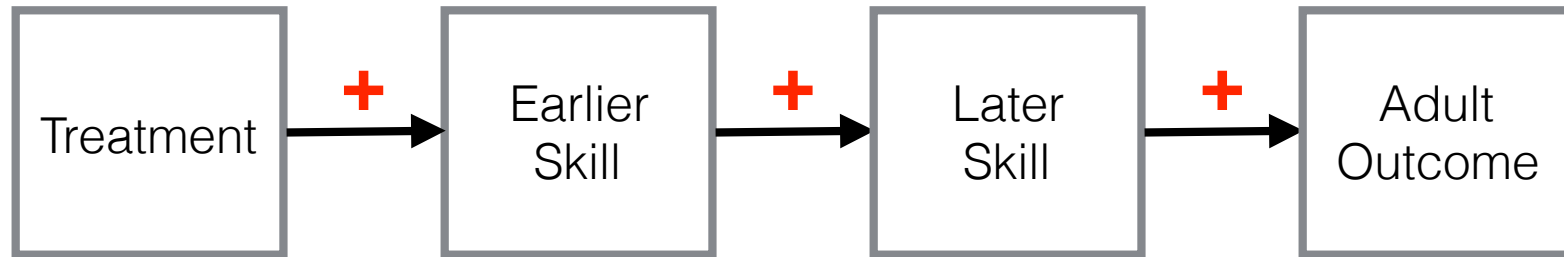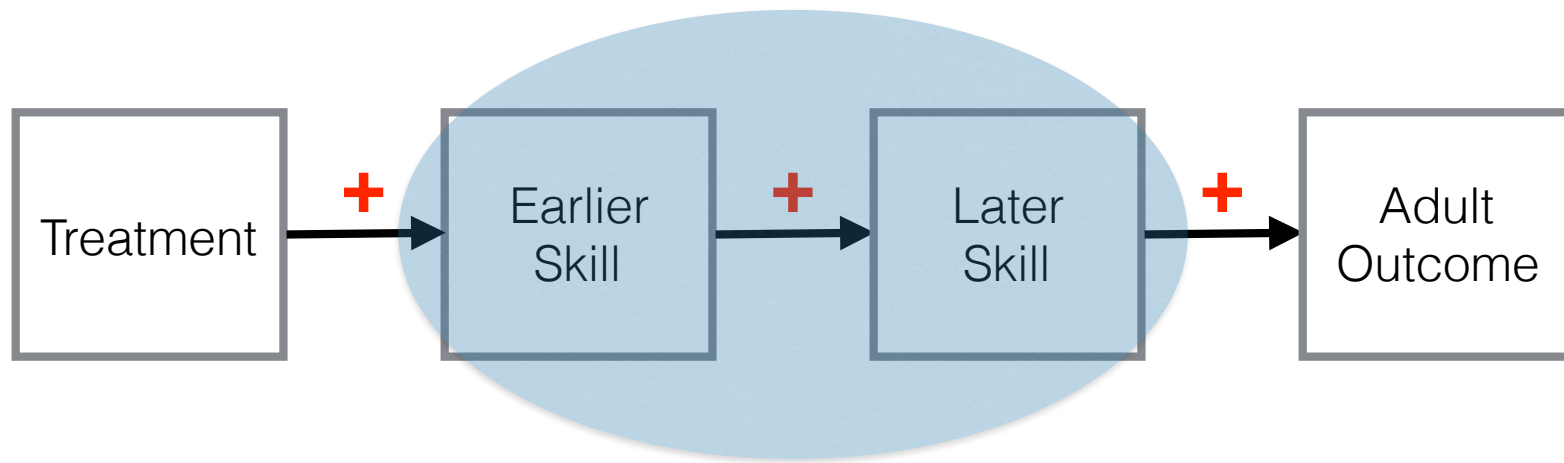# Early Academic Skill-Building: Developmental Processes and Implications for Intervention

Drew Bailey
School of Education
University of California, Irvine
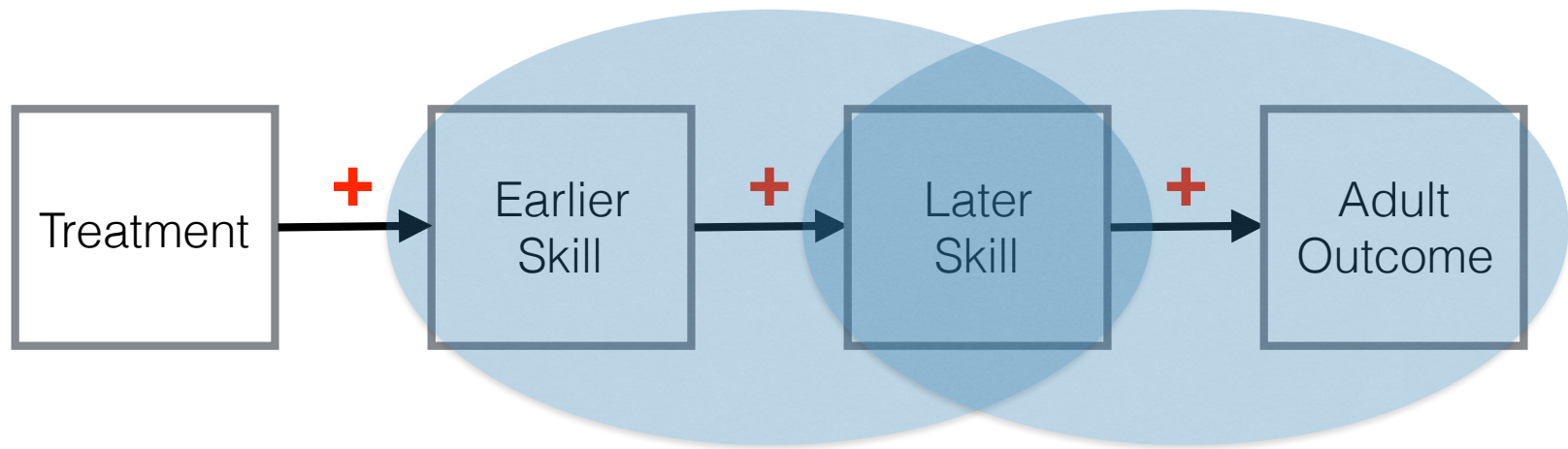
CCWD: Where I fit

```
┌──────────┐  +   ┌──────────┐  +   ┌──────────┐  +   ┌──────────┐
│          │─────▶│  Earlier │─────▶│   Later  │─────▶│   Adult  │
│ Treatment│      │   Skill  │      │   Skill  │      │  Outcome │
│          │      │          │      │          │      │          │
└──────────┘      └──────────┘      └──────────┘      └──────────┘
```

CCWD: Where I fit

```
Treatment  --+-->  Earlier Skill  --+-->  Later Skill  --+-->  Adult Outcome
```

CCWD: Where I fit

CCWD: Where I fit
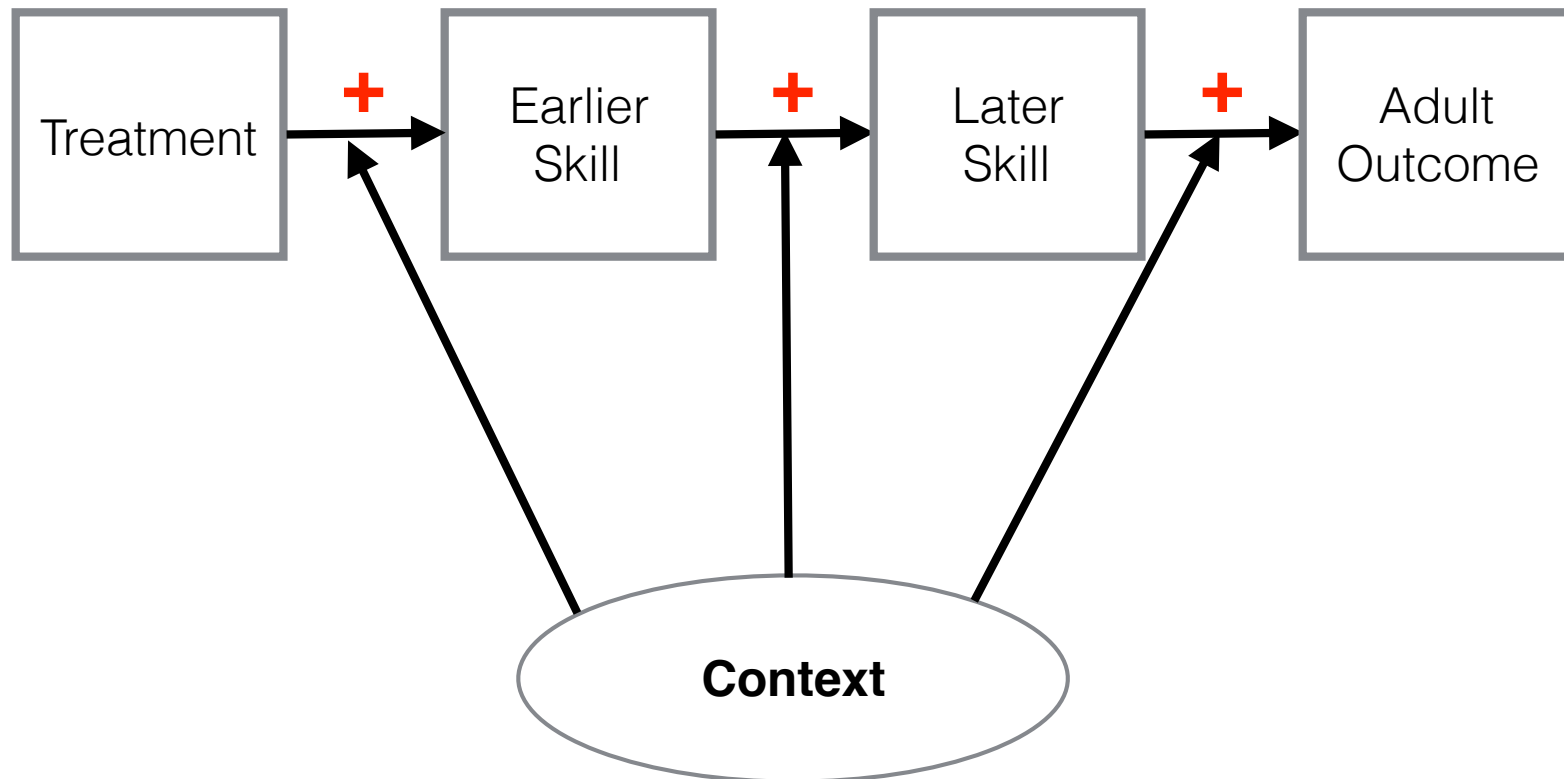
CCWD: ~~Where~~ **Do** I fit **?**
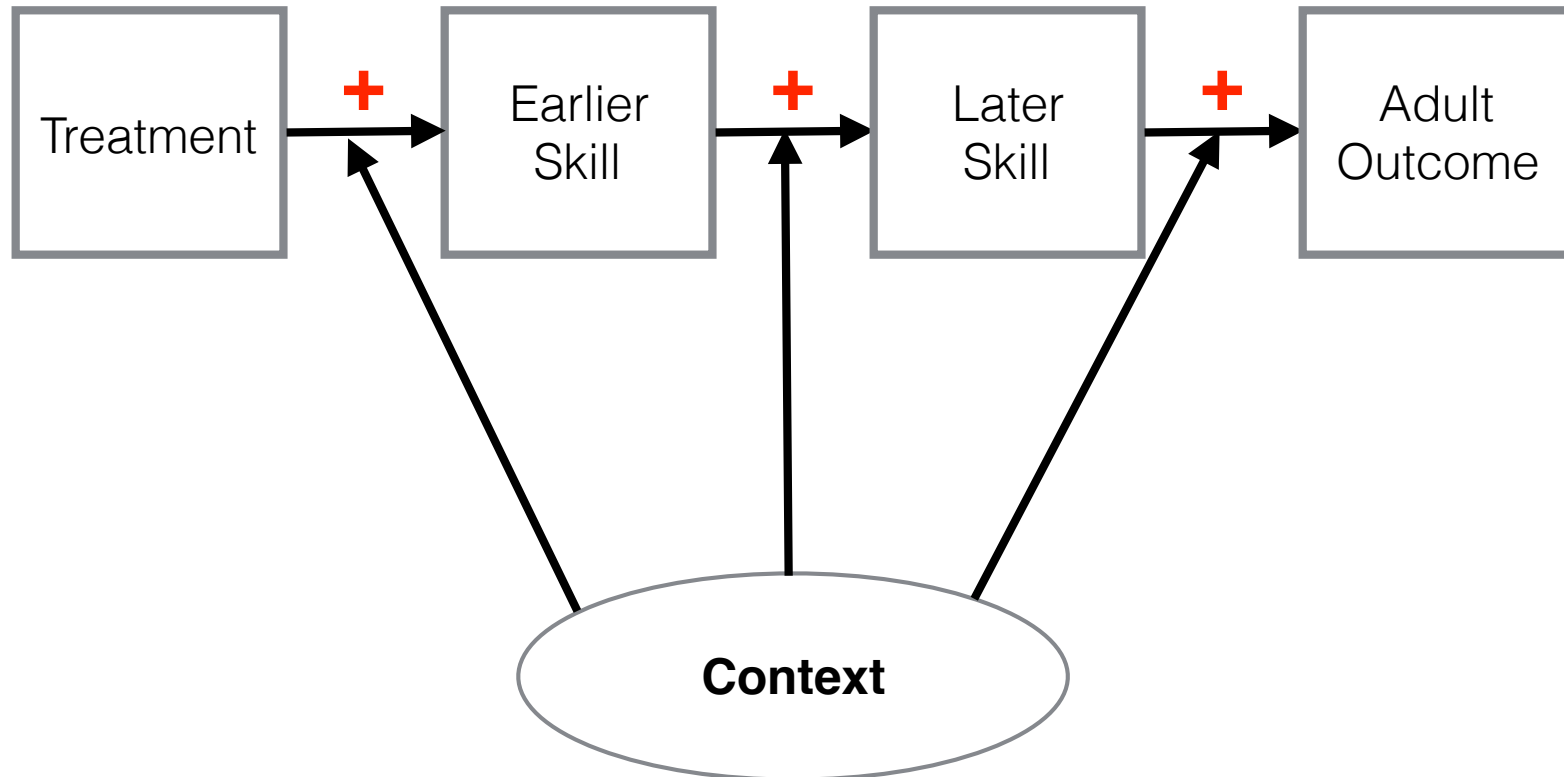
WARNING

Things I DON'T believe

Things I DON'T believe

- The factors producing stable individual differences in children's academic achievement are unmodifiable, in principle.

Things I DON'T believe

- The factors producing stable individual differences in children's academic achievement are unmodifiable, in principle.

- All early childhood education programs have no long-term benefits.

Things I believe

- The causal effects of one-time boosts to children's early academic skills on their much later academic skills are likely to be small.

Things I believe

- The causal effects of one-time boosts to children's early academic skills on their much later academic skills are likely to be small.

- If skill building is the mechanism we're/you're interested in, it might help to change some of our research practices and priorities.

Fadeout.

IQ impacts in Perry

End of program

Age
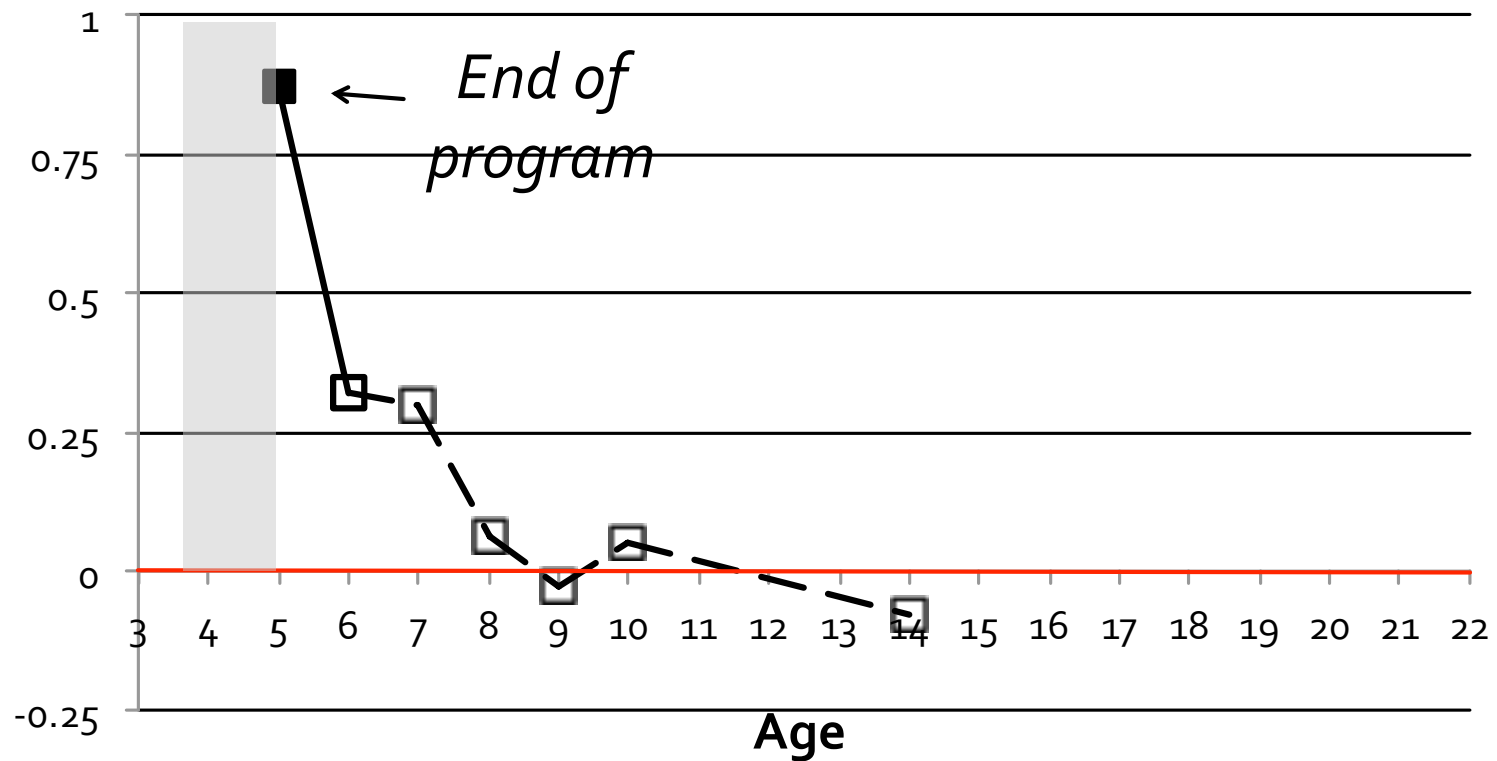
Solid marker denotes p<.05

Is it different for math?

Is it different for math?

- Theoretical distinction between achievement and aptitude
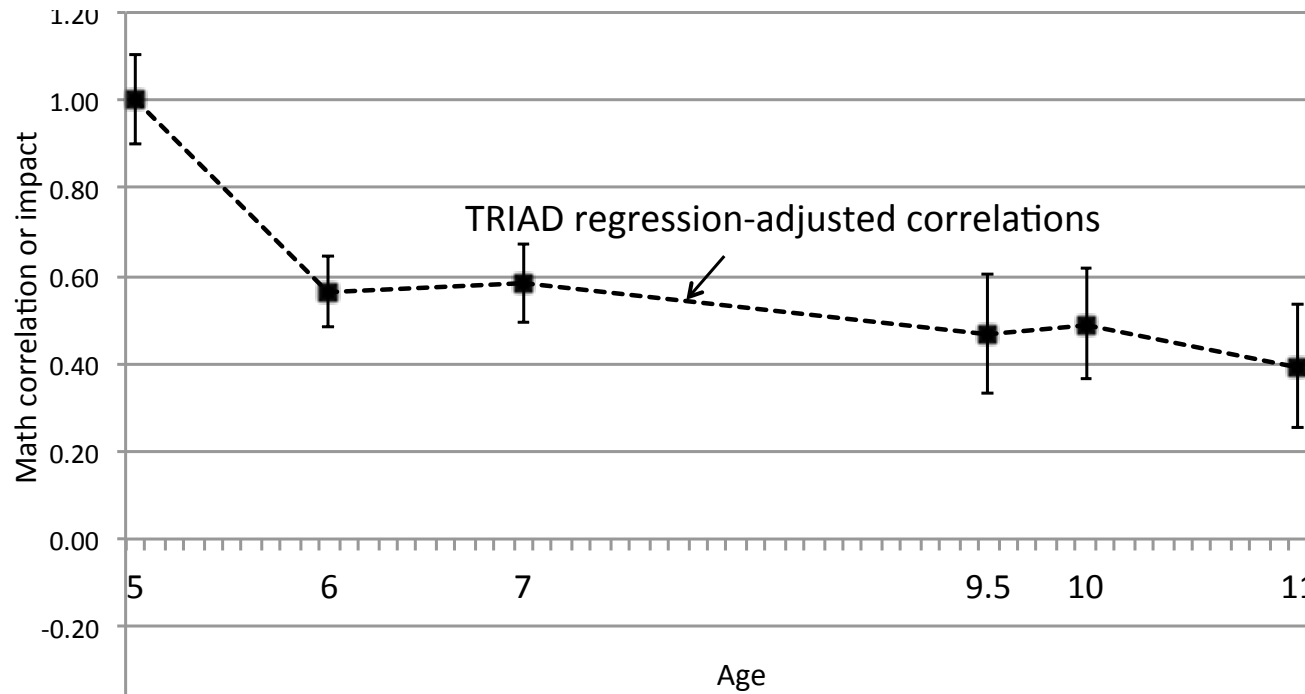
Is it different for math?

- Theoretical distinction between achievement and aptitude

- Clear vertical transfer in math learning (e.g., counting, addition, multiplication)

Is it different for math?

- Theoretical distinction between achievement and aptitude

- Clear vertical transfer in math learning (e.g., counting, addition, multiplication)

- Supportive correlational research

Is it different for math?

Controls: pre-k entry math, SES, ELL status, pre-k entry age

From Bailey, Duncan, Watts, Clements, & Sarama (2018, *American Psychologist*)
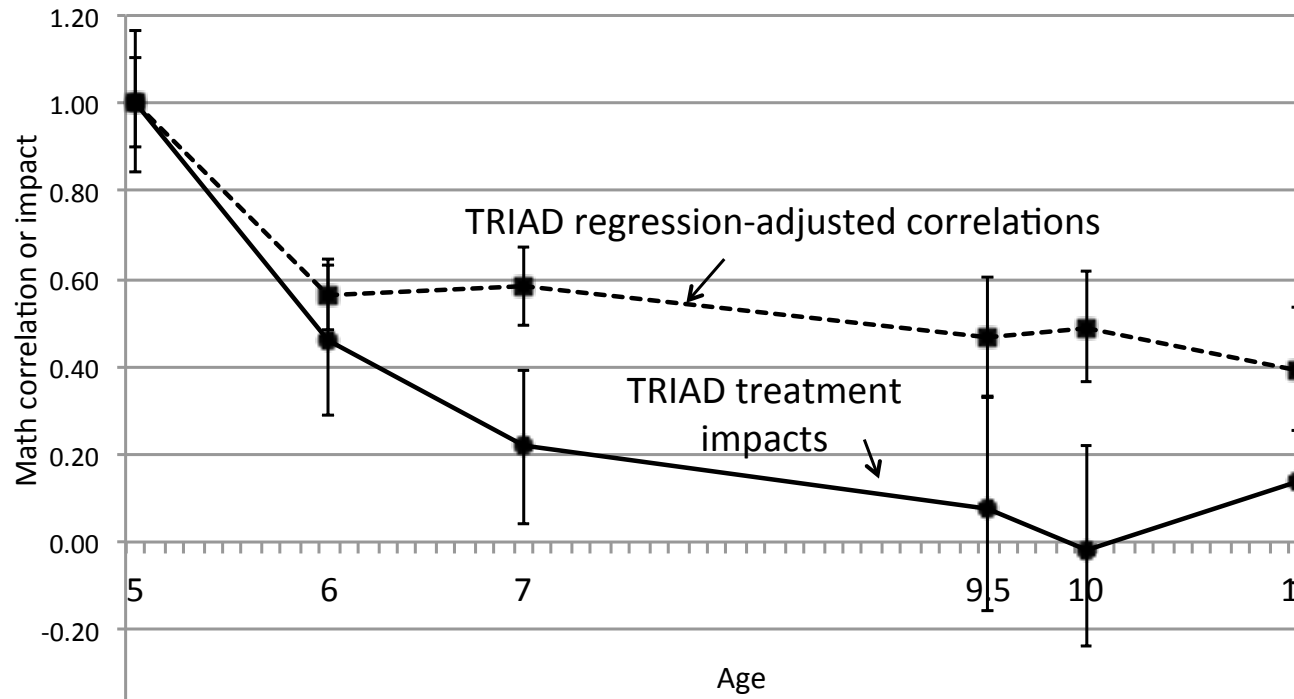
Is it different for math?

Controls: pre-k entry math, SES, ELL status, pre-k entry age

From Bailey, Duncan, Watts, Clements, & Sarama (2018, *American Psychologist*)

Controls:
IQ, working memory, reading achievement, SES, others

Earlier Mathematics Achievement

Later Mathematics Achievement

+
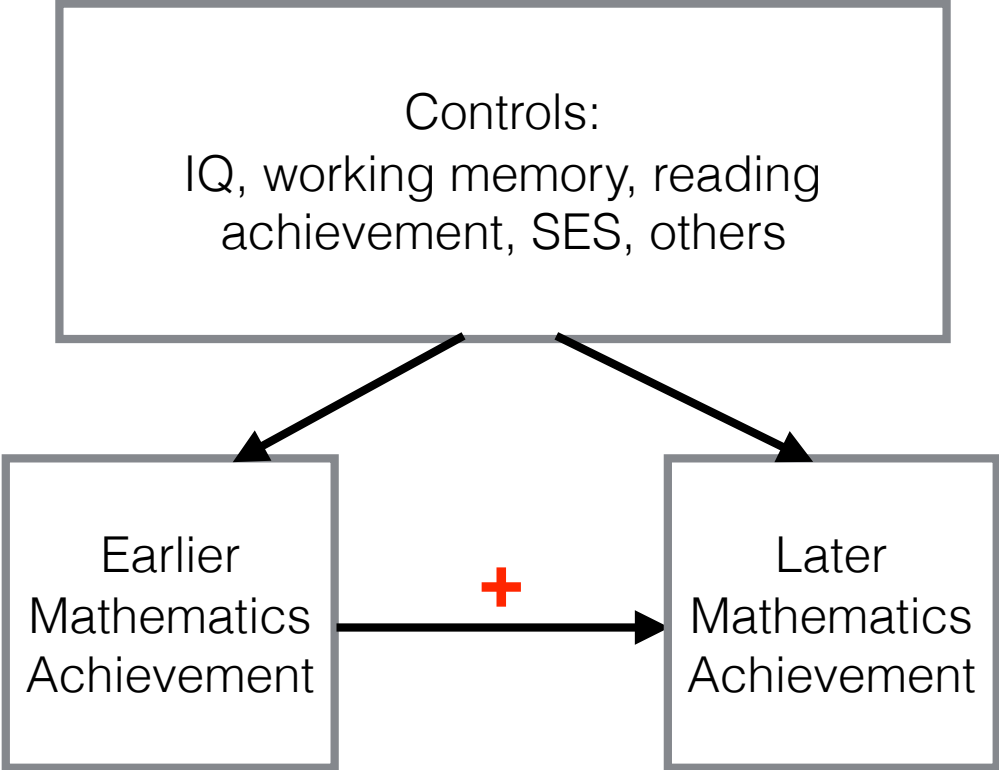
RESEARCH ARTICLE

Statistically Controlling for Confounding Constructs Is Harder than You Think

Jacob Westfall*, Tal Yarkoni
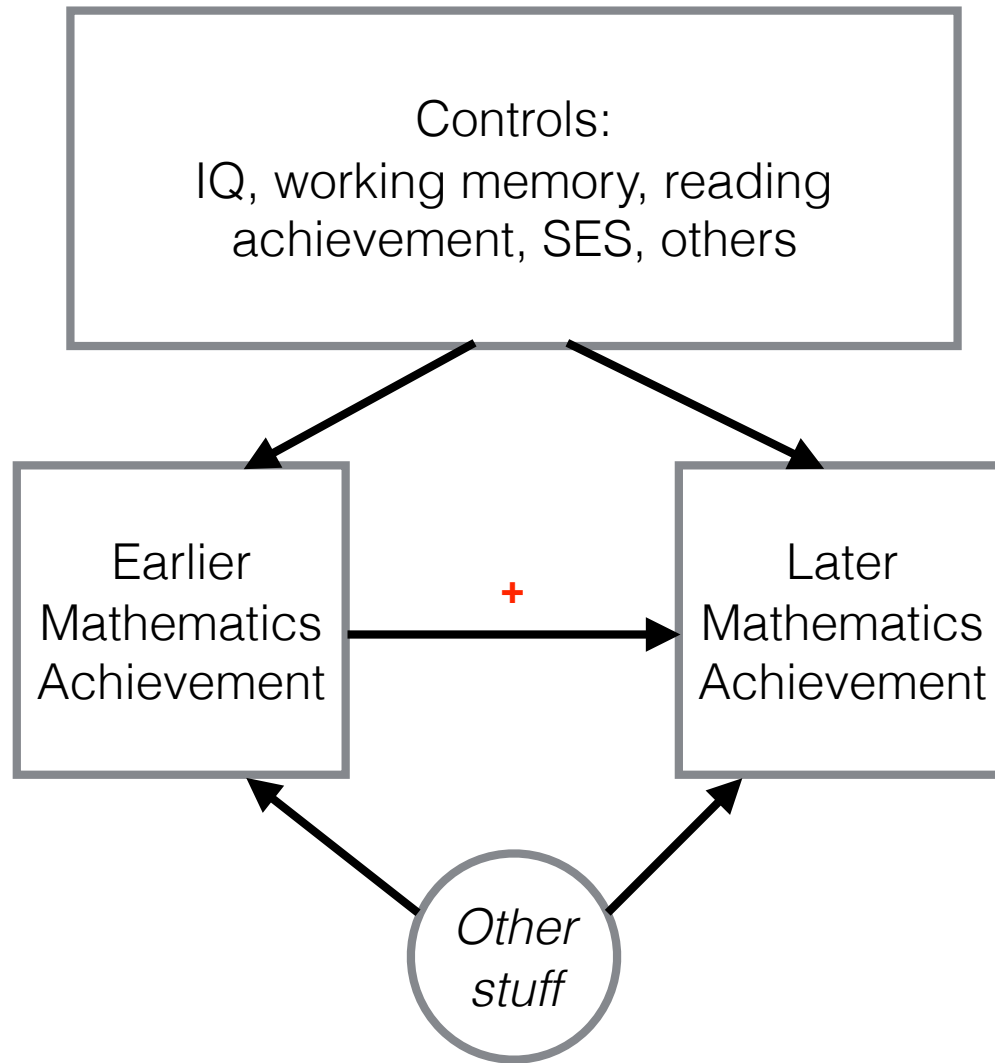
Boring, plausible alternative theory

Controls:
IQ, working memory, reading achievement, SES, others

Earlier Mathematics Achievement

+

Later Mathematics Achievement

vs.

Controls:
IQ, working memory, reading achievement, SES, others

Earlier Mathematics Achievement

+

Later Mathematics Achievement

Other stuff

Methods: can we do better than this?

Controls:
IQ, working memory, reading achievement, SES, others

Earlier Mathematics Achievement

+

Later Mathematics Achievement

vs.

Controls:
IQ, working memory, reading achievement, SES, others

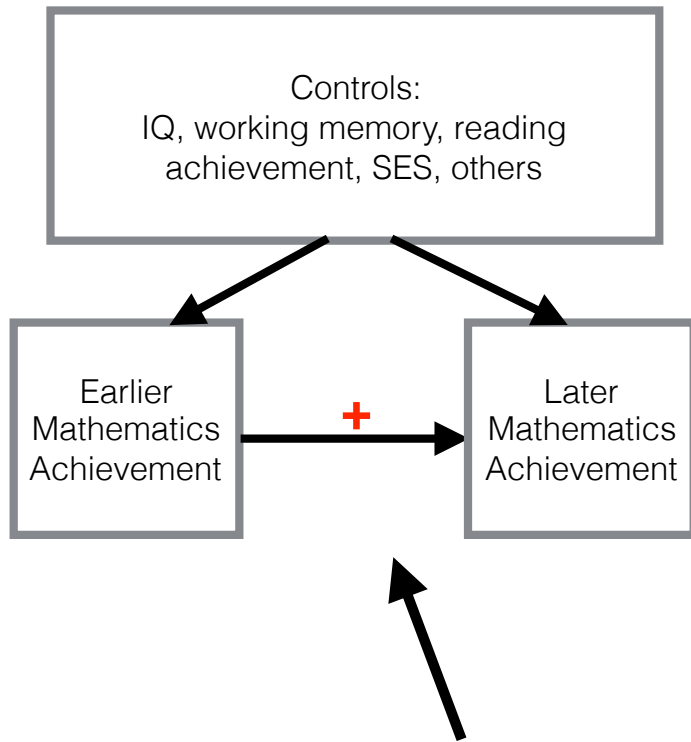Earlier Mathematics Achievement

+

Later Mathematics Achievement

Other stuff

Methods: can we do better than this?

i.e., how do we account for this?
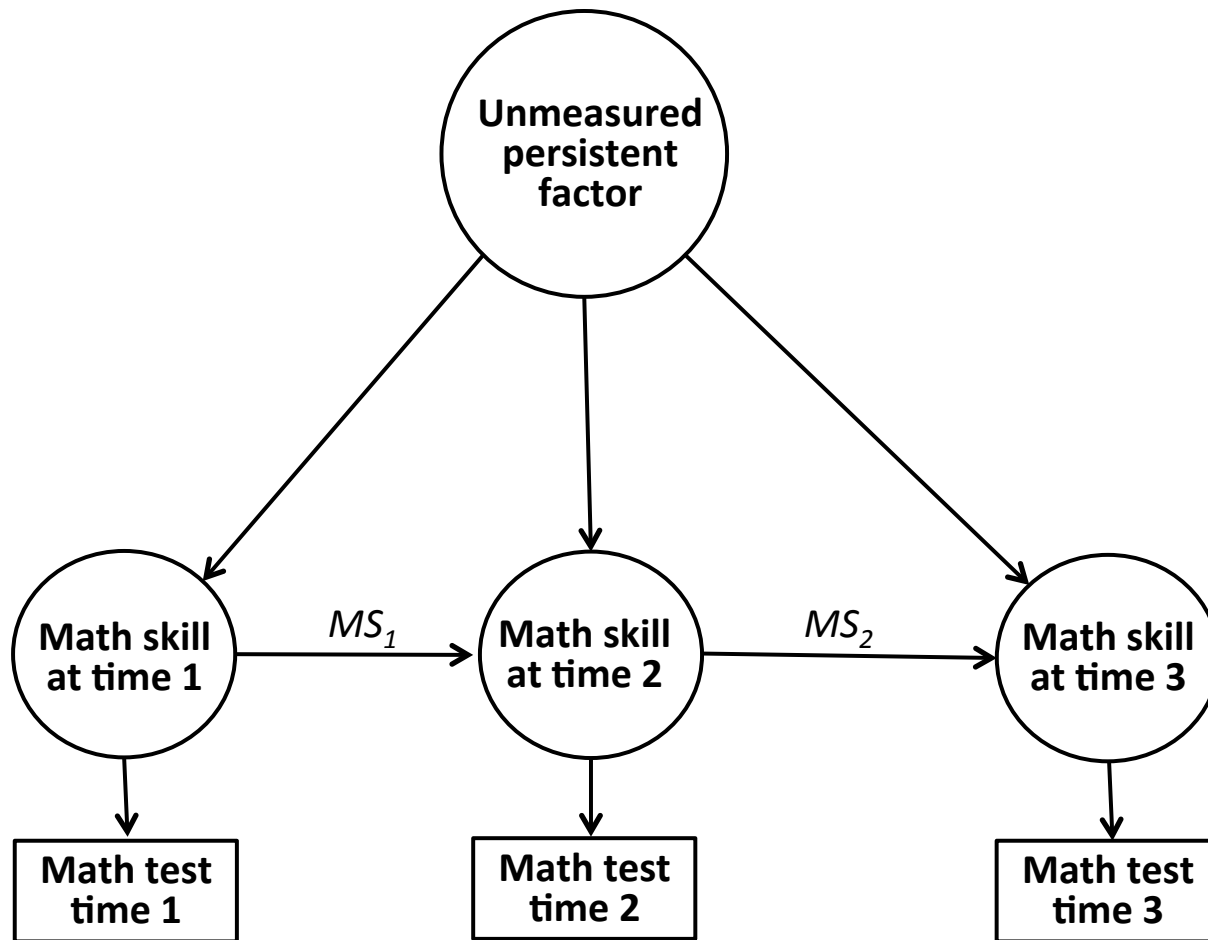
Latent state-trait model (Steyer, 1987)
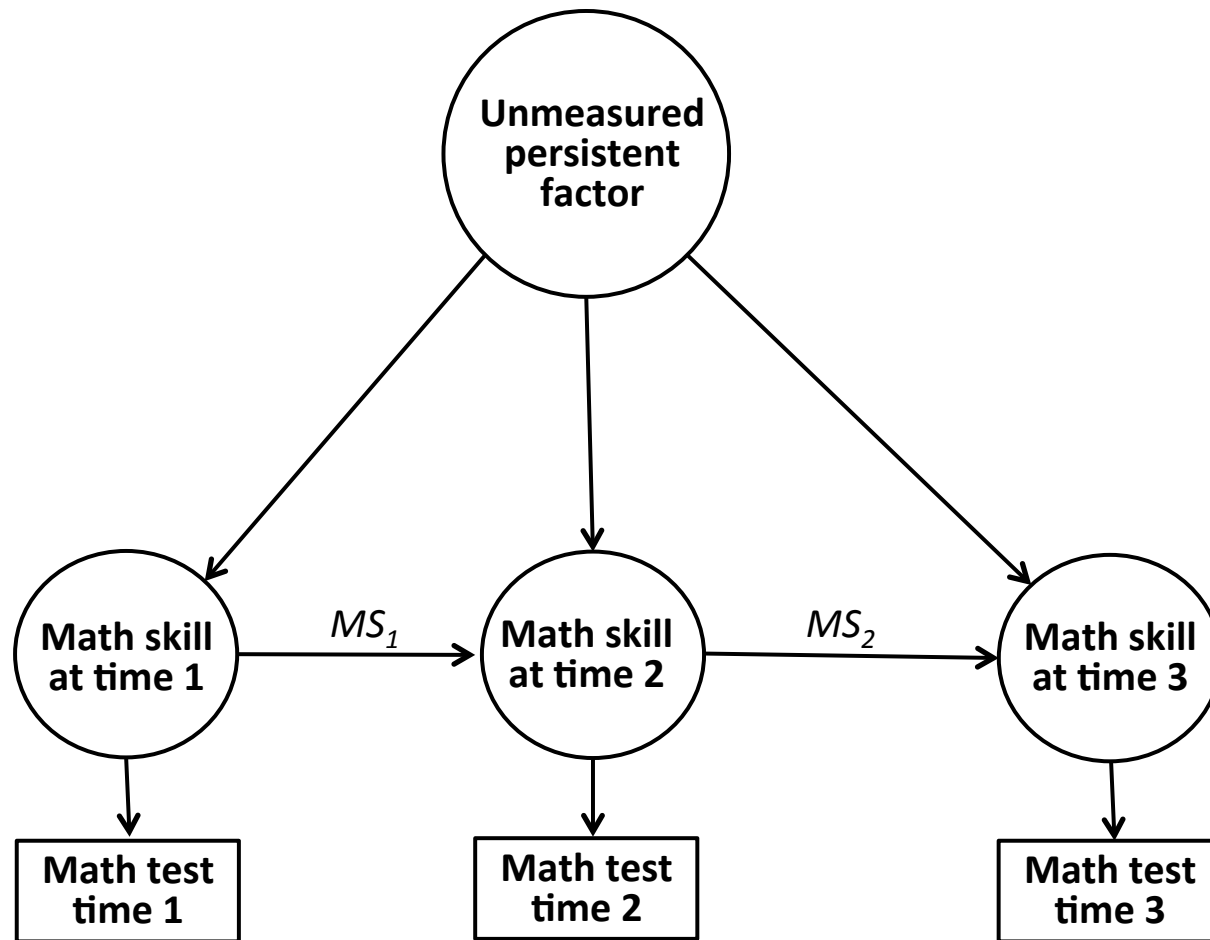
Average 1-year *MS* estimate from 3 datasets: **.35**

From Bailey, Watts, Littlefield, & Geary (2014; *Psych Science*);
Bailey et al., (2018, *American Psychologist*)
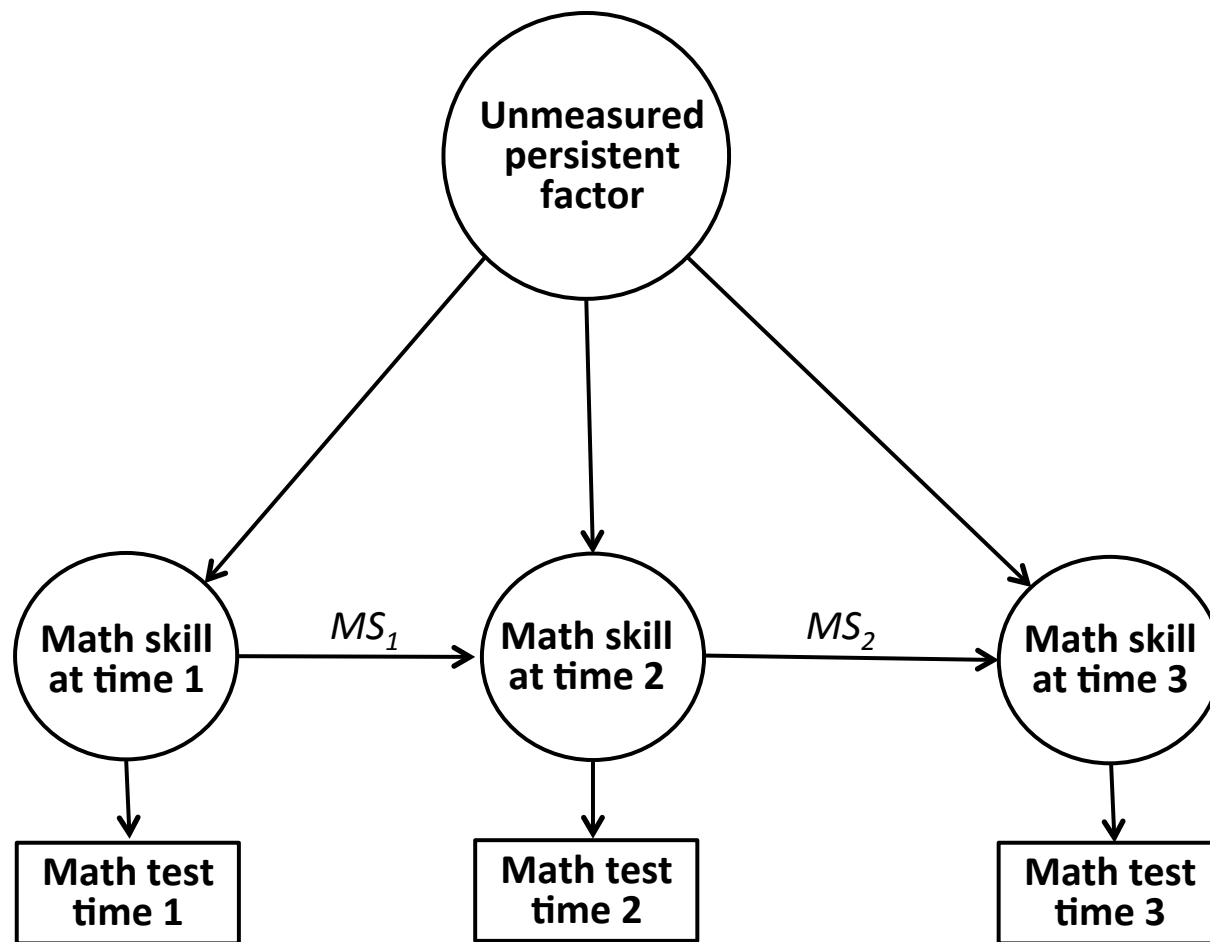
Average 1-year *MS* estimate from 3 datasets: **.35**

For meta-analytic AR estimates for personality, see Anusic & Schimmack (2016, *JPSP*)

From Bailey, Watts, Littlefield, & Geary (2014; *Psych Science*);
Bailey et al., (2018, *American Psychologist*)

Figure 5: Correlations inferred from *MS* path estimates in Table 1
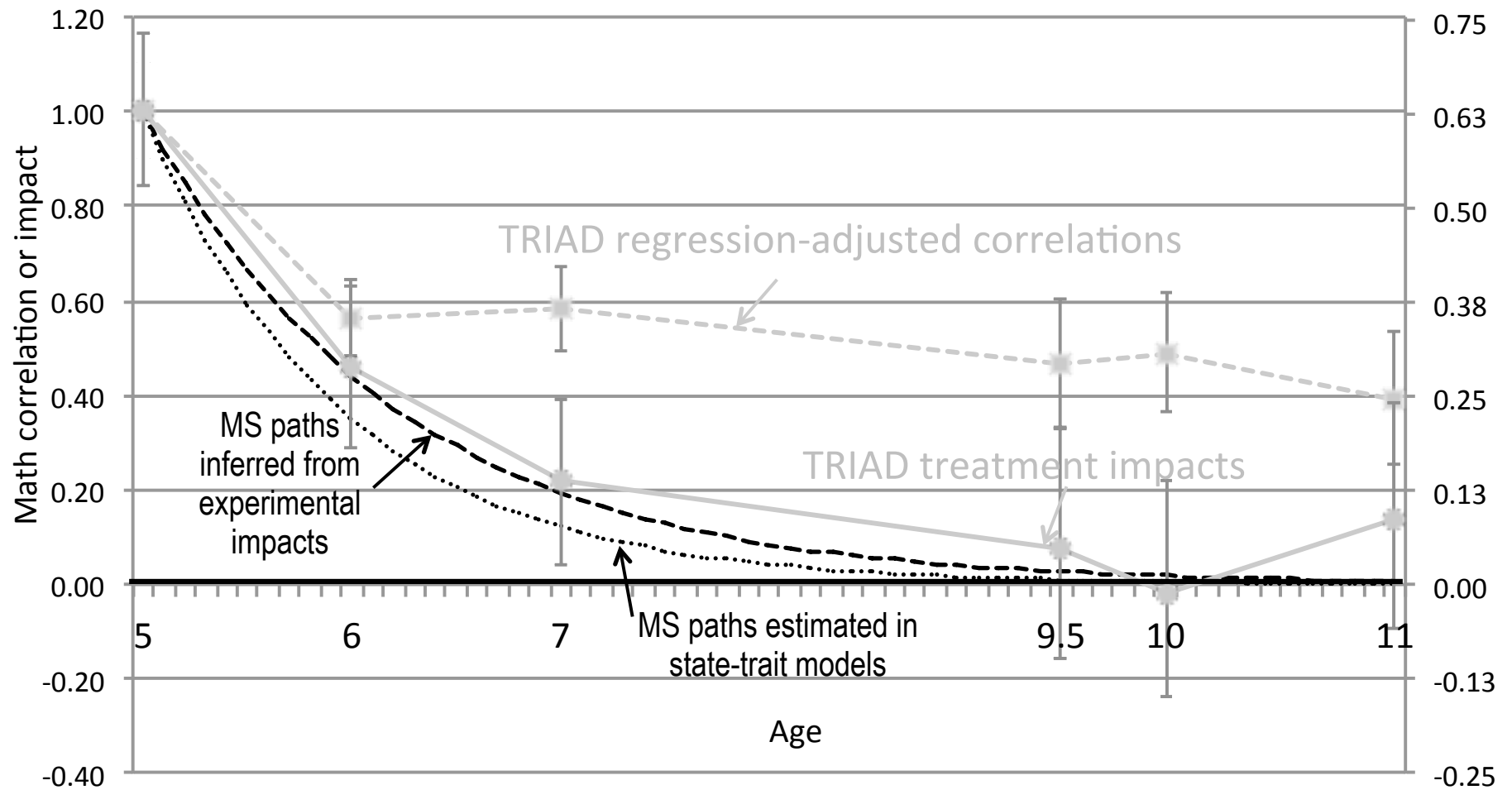
Figure 2: Cognitive impacts in 67 ECE studies
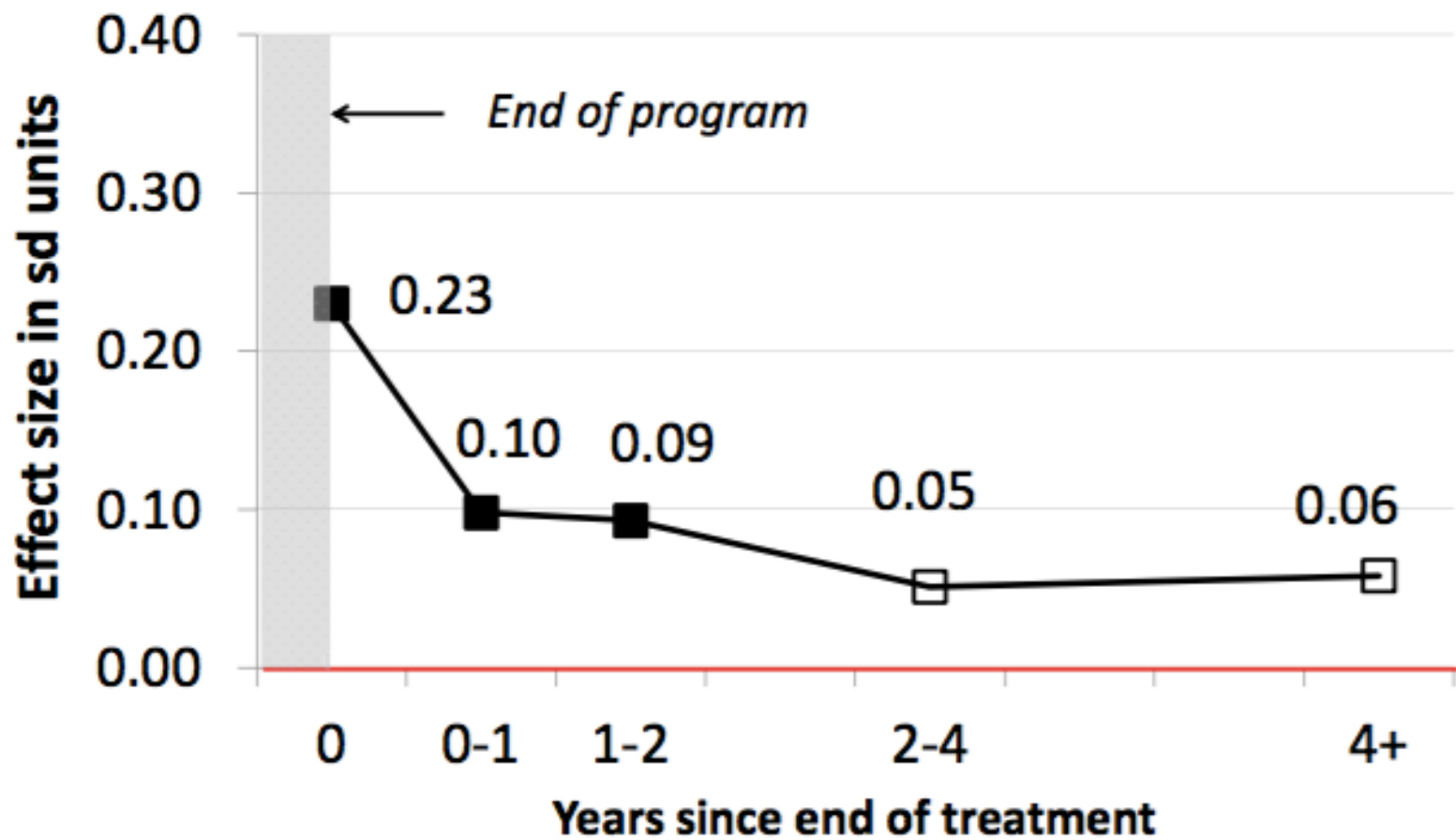
Data from Li et al. (2017)

Fadeout/catchup
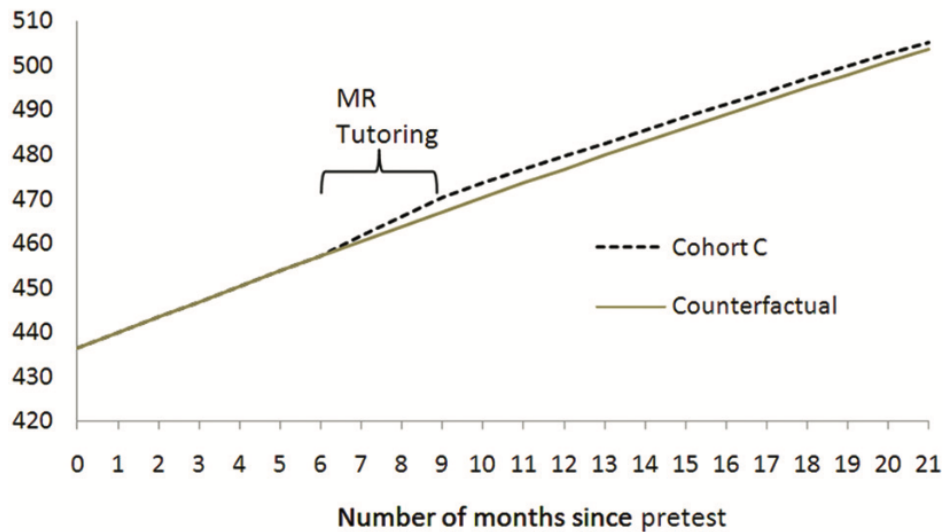
From Smith, Cobb, Farran,
Cordray, & Munter
(2013, *AERJ*)

Fadeout/catchup
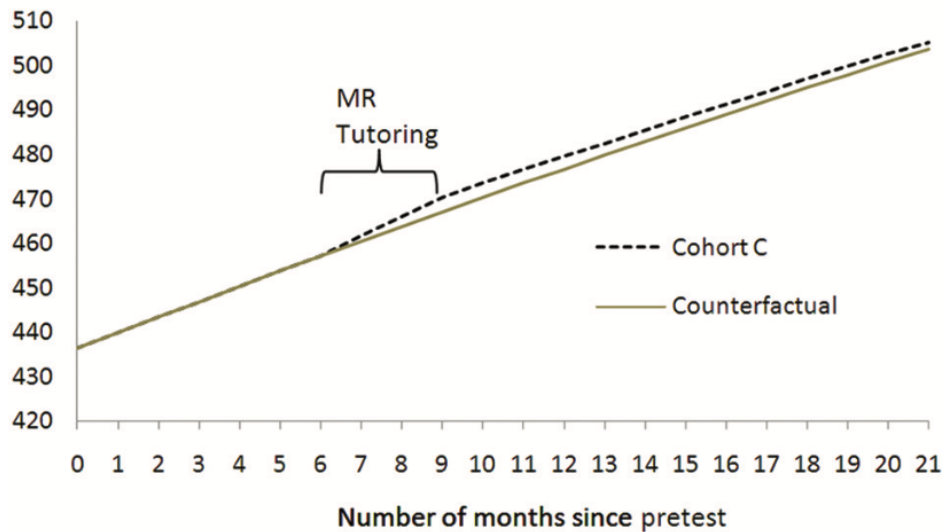
From Smith, Cobb, Farran, Cordray, & Munter (2013, *AERJ*)
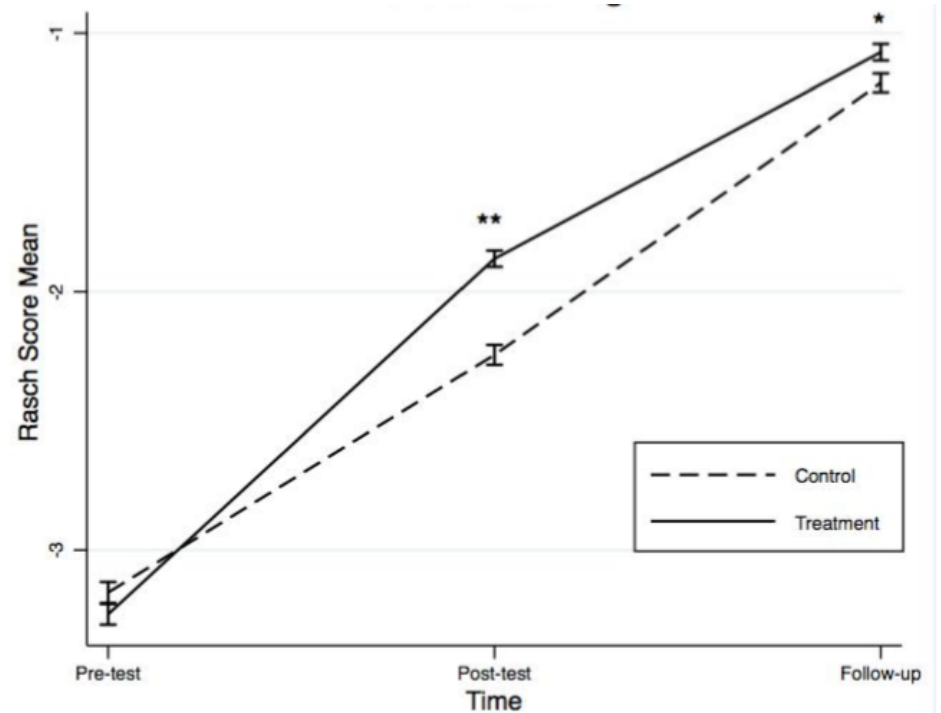
From Bailey, Nguyen, Jenkins, Domina, Clements, & Sarama (2016, *Developmental Psychology*)

What's going on?

Constraining Content
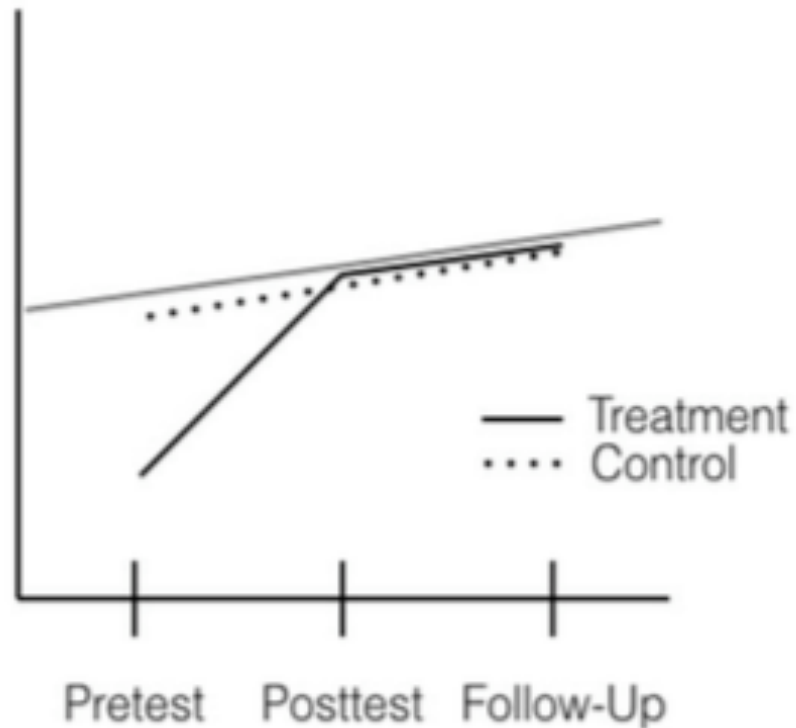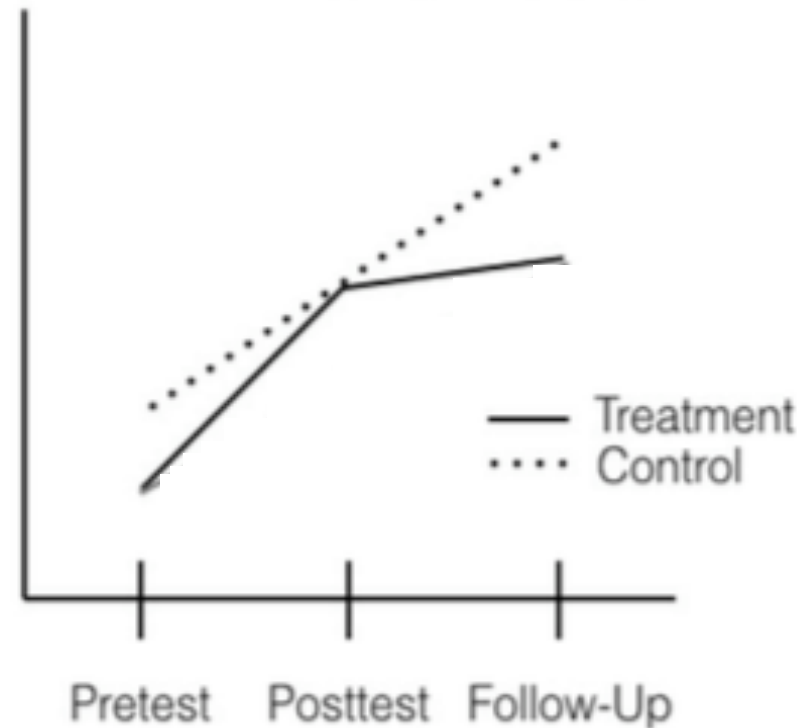
Treatment
Control

Pretest   Posttest   Follow-Up

From Bailey, Nguyen, Jenkins, Domina, Clements, & Sarama
(2016, *Developmental Psychology*)

What's going on?

From Bailey, Nguyen, Jenkins, Domina, Clements, & Sarama
(2016, *Developmental Psychology*)

What's going on?
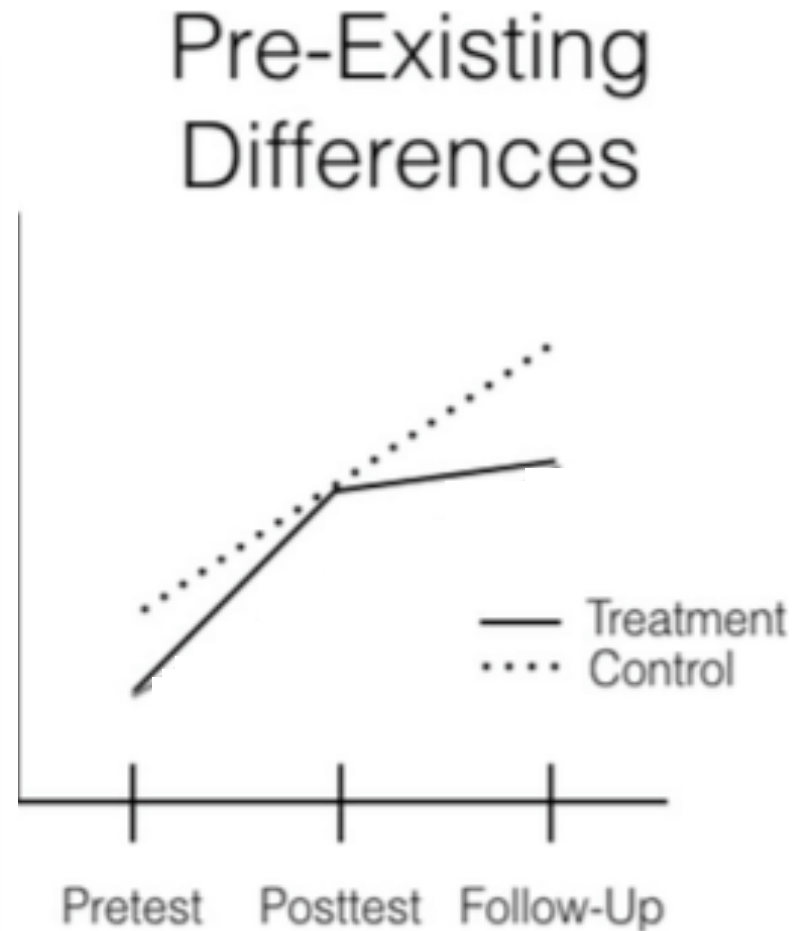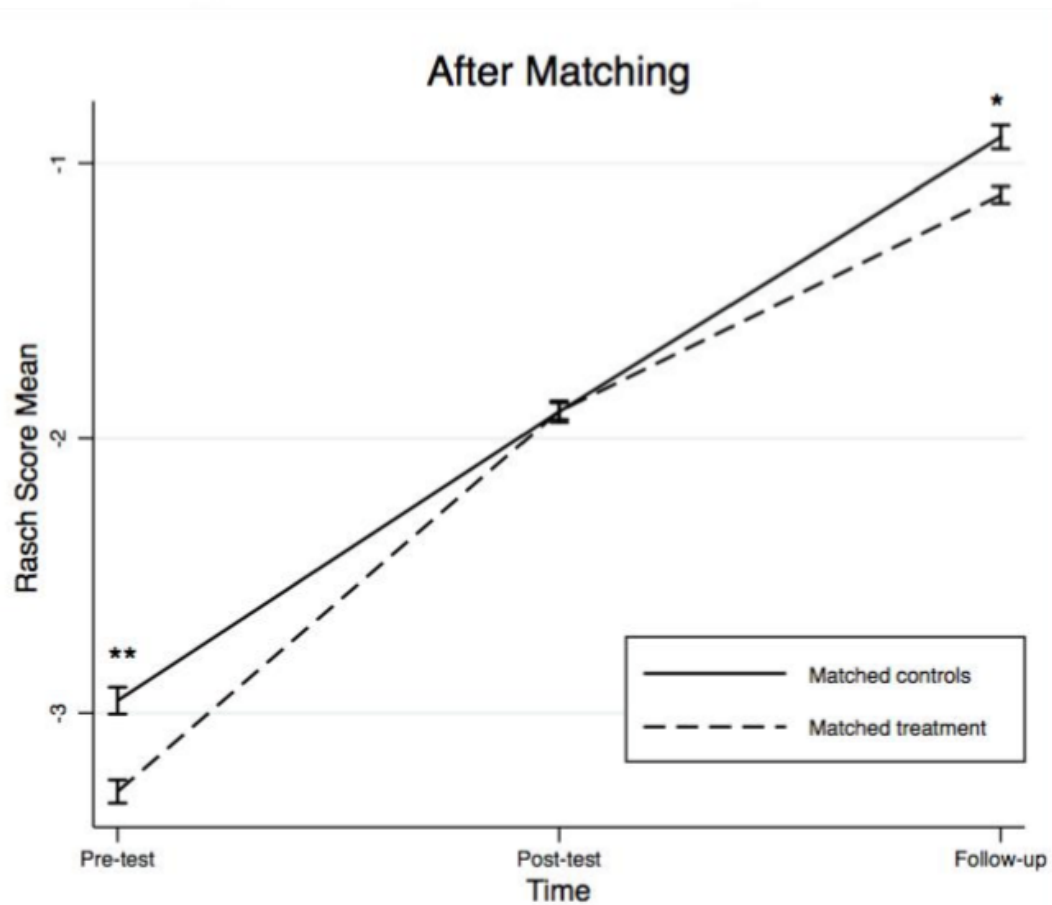
From Bailey, Nguyen, Jenkins, Domina, Clements, & Sarama
(2016, *Developmental Psychology*)

But why the classic ECE findings?

But why the classic ECE findings?

- Possible explanations

  - The "right" kinds of skills

The "right" kinds of skills?

The "right" kinds of skills?

- Can we identify skills that are:

  - Malleable through intervention,

The "right" kinds of skills?

- Can we identify skills that are:

  - Malleable through intervention,

  - Fundamental for success,

The "right" kinds of skills?

- Can we identify skills that are:

  - Malleable through intervention,

  - Fundamental for success,

  - And would not develop quickly in most counterfactual conditions?

From Bailey, Duncan, Odgers, & Yu
(2017, *Journal of Research on Educational Effectiveness*)

Average Annual Gain in Effect Size From Nationally Normed Tests

Grade transition	Reading tests		Math tests	
	Mean	Margin of error	Mean	Margin of error
Grade K–1	1.52	±0.21	1.14	±0.49
Grade 1–2	0.97	±0.10	1.03	±0.14

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)

Average Annual Gain in Effect Size From Nationally Normed Tests

Grade transition	Reading tests		Math tests	
	Mean	Margin of error	Mean	Margin of error
Grade K–1	1.52	±0.21	1.14	±0.49
Grade 1–2	0.97	±0.10	1.03	±0.14
Grade 2–3	0.60	±0.10	0.89	±0.16
Grade 3–4	0.36	±0.12	0.52	±0.14
Grade 4–5	0.40	±0.06	0.56	±0.11

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)

Average Annual Gain in Effect Size From Nationally Normed Tests

Grade transition	Reading tests		Math tests	
	Mean	Margin of error	Mean	Margin of error
Grade K–1	1.52	±0.21	1.14	±0.49
Grade 1–2	0.97	±0.10	1.03	±0.14
Grade 2–3	0.60	±0.10	0.89	±0.16
Grade 3–4	0.36	±0.12	0.52	±0.14
Grade 4–5	0.40	±0.06	0.56	±0.11
Grade 5–6	0.32	±0.11	0.41	±0.08
Grade 6–7	0.23	±0.11	0.30	±0.06
Grade 7–8	0.26	±0.03	0.32	±0.05

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)

Average Annual Gain in Effect Size From Nationally Normed Tests

Grade transition	Reading tests		Math tests	
	Mean	Margin of error	Mean	Margin of error
Grade K–1	1.52	±0.21	1.14	±0.49
Grade 1–2	0.97	±0.10	1.03	±0.14
Grade 2–3	0.60	±0.10	0.89	±0.16
Grade 3–4	0.36	±0.12	0.52	±0.14
Grade 4–5	0.40	±0.06	0.56	±0.11
Grade 5–6	0.32	±0.11	0.41	±0.08
Grade 6–7	0.23	±0.11	0.30	±0.06
Grade 7–8	0.26	±0.03	0.32	±0.05
Grade 8–9	0.24	±0.10	0.22	±0.10
Grade 9–10	0.19	±0.08	0.25	±0.07
Grade 10–11	0.19	±0.17	0.14	±0.16
Grade 11–12	0.06	±0.11	0.01	±0.14

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)

Average Annual Gain in Effect Size From Nationally Normed Tests

WARNING

Grade transition	Reading tests		Math tests	
	Mean	Margin of error	Mean	Margin of error
Grade K–1	1.52	±0.21	1.14	±0.49
Grade 1–2	0.97	±0.10	1.03	±0.14
Grade 2–3	0.60	±0.10	0.89	±0.16
Grade 3–4	0.36	±0.12	0.52	±0.14
Grade 4–5	0.40	±0.06	0.56	±0.11
Grade 5–6	0.32	±0.11	0.41	±0.08
Grade 6–7	0.23	±0.11	0.30	±0.06
Grade 7–8	0.26	±0.03	0.32	±0.05
Grade 8–9	0.24	±0.10	0.22	±0.10
Grade 9–10	0.19	±0.08	0.25	±0.07
Grade 10–11	0.19	±0.17	0.14	±0.16
Grade 11–12	0.06	±0.11	0.01	±0.14

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)

The "right" kinds of skills?

- Can we identify skills that are:

  - Malleable through intervention,

  - Fundamental for success,

  - And would not develop quickly in most counterfactual conditions?

Why is this list so tragically short?

- Tradeoffs between *trifecta* criteria:

  - Fundamental AND malleable (e.g., basic language and literacy) are already aggressively targeted, and therefore likely to develop under counterfactual conditions.

Why is this list so tragically short?

- Tradeoffs between *trifecta* criteria:

  - Fundamental AND malleable (e.g., basic language and literacy) are already aggressively targeted, and therefore likely to develop under counterfactual conditions.

  - Clearly fundamental skills that do not develop under most counterfactual conditions are not likely to be malleable by scalable interventions (perhaps too pessimistic).

Why is this list so tragically short?

- Tradeoffs between *trifecta* criteria:

  - Fundamental AND malleable (e.g., basic language and literacy) are already aggressively targeted, and therefore likely to develop under counterfactual conditions.
  WARNING
  - Clearly fundamental skills that do not develop under most counterfactual conditions are not likely to be malleable by scalable interventions (perhaps too pessimistic).

  WARNING

But why the classic ECE findings?

- Possible explanations

  - The "right" kinds of skills

From Bailey, Duncan, Odgers, & Yu
(2017, *Journal of Research on Educational Effectiveness*)

But why the classic ECE findings?

- Possible explanations

  - The "right" kinds of skills

  - "Foot-in-the-Door" pathways: The right affordances at the right times get children through a period of risk

Possible *foot in the door* pathways

- Non-trifecta skills that keep children from being retained in school, kicked out of school, choosing a bad peer group

Possible *foot in the door* pathways

- Non-trifecta skills that keep children from being retained in school, kicked out of school, choosing a bad peer group

- Suggestive evidence from Chicago Double Dose Algebra evaluation

Possible *foot in the door* pathways

Can we avoid negative and promote positive developmental cascades?

A cascade model of Dodge et al (2008):

Adverse early context

Early harsh parenting

Poor school readiness

Conduct problems

School failure

Low parent monitoring

Deviant Peers

Adolescent violence

Early life – *Adolescence*

Possible *foot in the door* pathways

Can we avoid negative and promote positive developmental cascades?

A cascade model of Dodge et al (2008):

```
Adverse                                    Adolescent
early                                       violence
context
        \                                          \
         Early harsh                        Deviant
         parenting                          Peers
                  \                        /
                  Poor school     Low parent
                  readiness       monitoring
                          \      /
                          Conduct
                          problems
                                \
                                School
                                failure
```
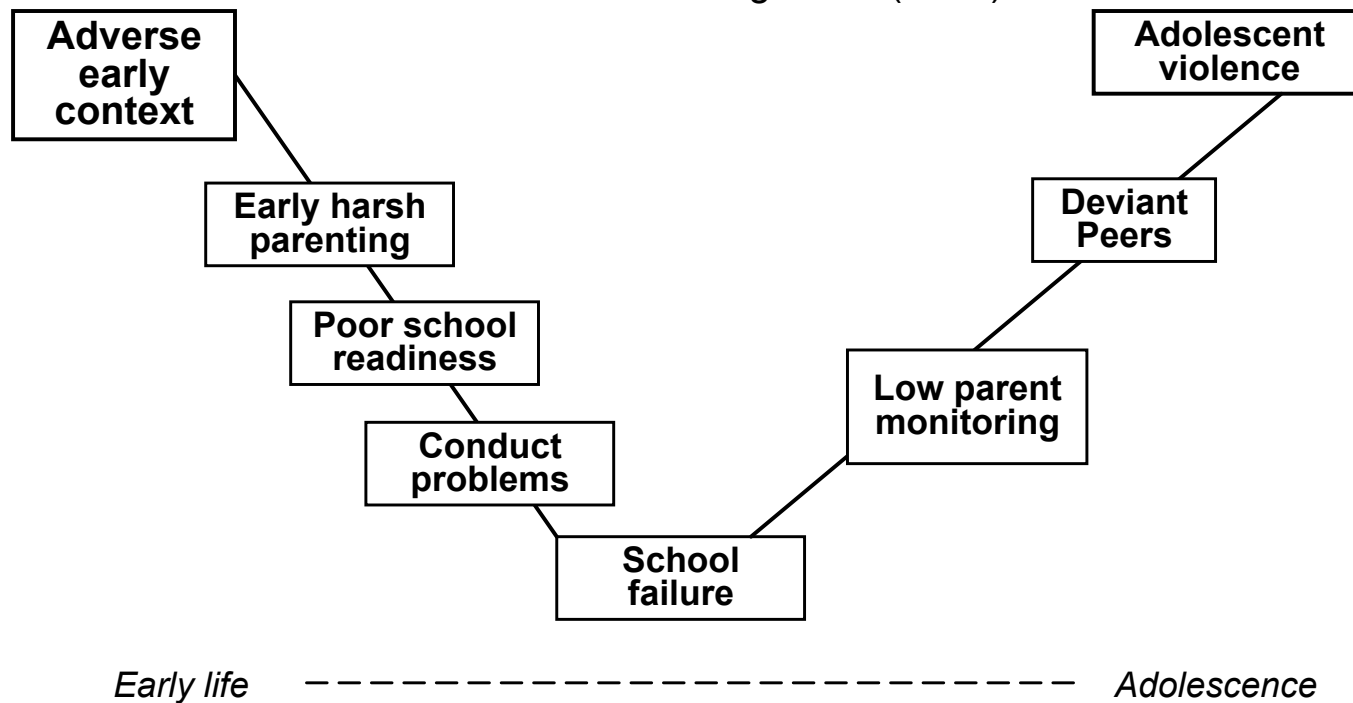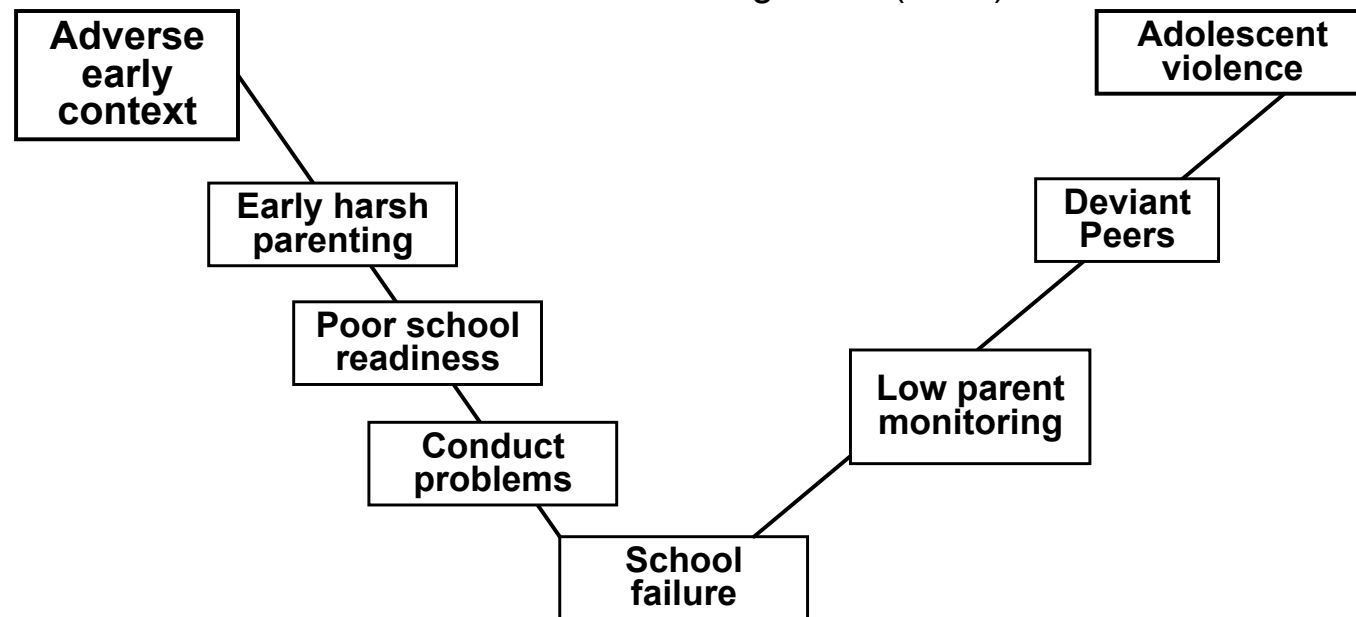
Early life – *Adolescence*

Problematic if these probabilities multiply.

Can we avoid negative and promote positive developmental cascades?

A cascade model of Dodge et al (2008):

Adverse early context

Early harsh parenting

Poor school readiness

Conduct problems

School failure

Low parent monitoring

Deviant Peers

Adolescent violence

Early life – *Adolescence*
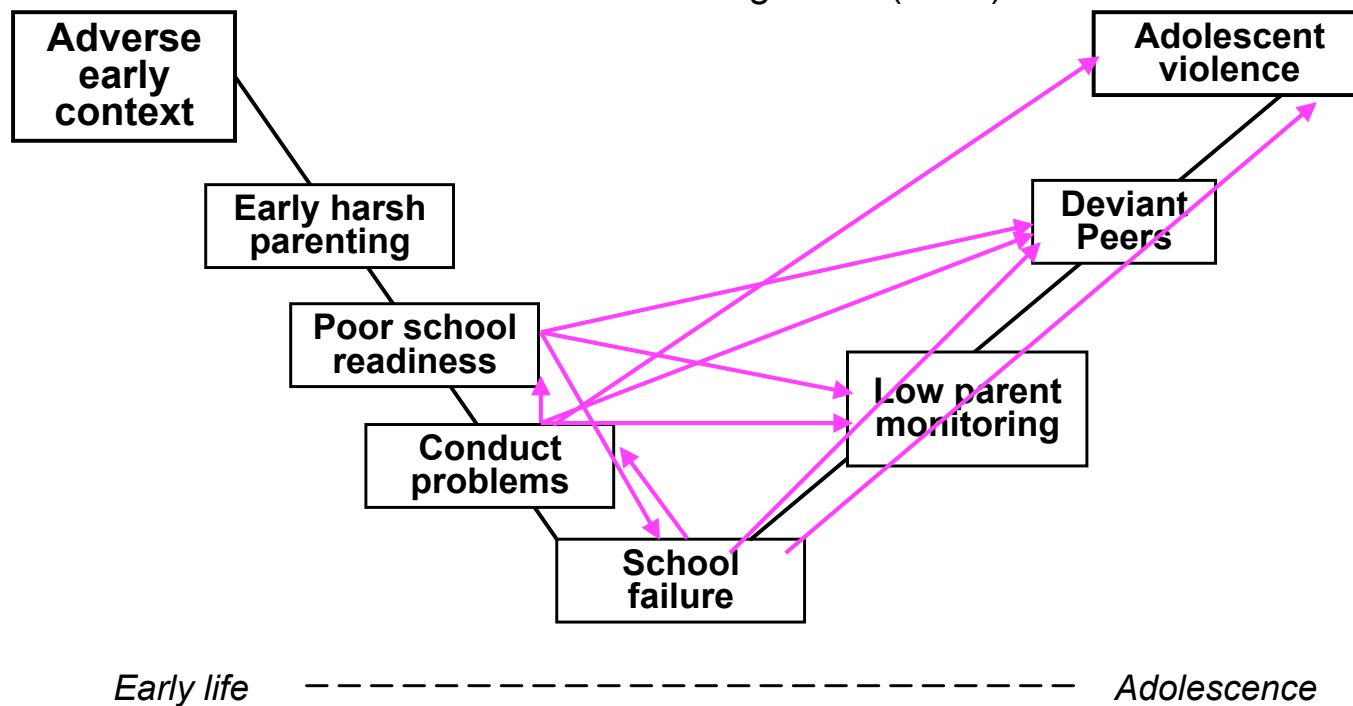
But, if temporary boosts increase the likelihood of thousands of foot in the door pathways, ...

Implications for Research

- Clarity of theory and methods

- More causally informative analysis

Implications for Research

- Clarity of theory and methods

- More causally informative analysis (but maybe I am "preaching to the choir")

Implications for Research

- Clarity of theory and methods

- More causally informative analysis by skill building researchers **+ Introduction to modern measurement theory for policy researchers**
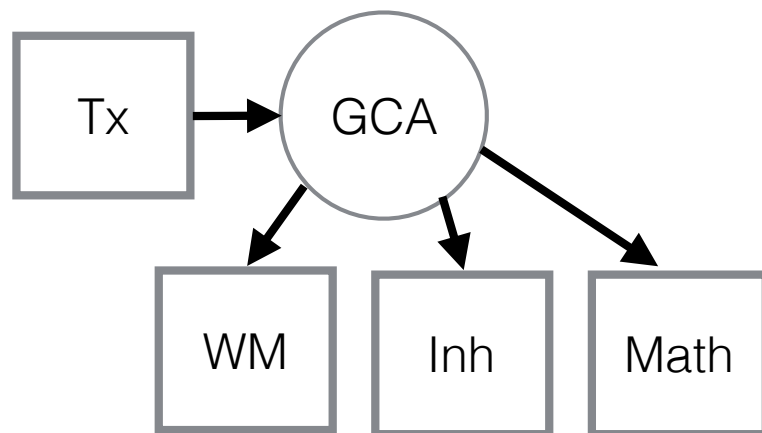
Implications for Research

- Clarity of theory and methods

- More causally informative analysis in personality research + Introduction to modern measurement theory for policy researchers

- Policy relevant field experimentation

Implications for Research

- Clarity of theory and methods

- More causally informative analysis in personality research + Introduction to modern measurement theory for policy researchers

- Policy relevant field experimentation

  - Follow-up data

Implications for Research

- Clarity of theory and methods

- More causally informative analysis in personality research + Introduction to modern measurement theory for policy researchers

- Policy relevant field experimentation

    - Follow-up data

    - Follow-up interventions

Acknowledgments

- TRIAD districts, teachers, & students

- Greg Duncan, Tyler Watts, Doug Clements, Julie Sarama, Tutrang Nguyen

- Dave Geary, Bob Siegler, Andrew Littlefield

- CCWD

Questions?

dhbailey@uci.edu

Perry Preschool

# of misdemeanor violent crimes, age 27 (-0.423**)	0.099 — 0.305
# of felony arrests, age 27 (-0.269**)	0.120 — 0.319
Jobless for more than 1 year, age 27 (-0.292*)	0.497 — 0.127
Ever tried drugs other than alcohol or weed, age 27 (-0.227**)	0.199 — 0.228 — 0.150
# of misdemeanor violent crimes, age 40 (-0.537**)	0.066 — 0.371
# of felony arrests, age 40 (-0.383**)	0.050 — 0.369
# of lifetime violent crimes, age 40 (-0.574**)	0.046 — 0.320
Months in all marriages, age 40 (39.6*)	0.185 — 0.224 — 0.269 — 0.352

Percent: 0 20 40 60 80 100

Legend: ▨ Cognitive Factor ■ Externalizing Behavior ◪ Academic Motivation ☐ Other Factors

Perry Preschool

# of misdemeanor violent crimes, age 27 (-0.423**)	0.099 ... 0.305
# of felony arrests, age 27 (-0.269**)	0.120 ... 0.319
Jobless for more than 1 year, age 27 (-0.292*)	0.497 ... 0.127
Ever tried drugs other than alcohol or weed, age 27 (-0.227**)	0.199 ... 0.228 ... 0.150
# of misdemeanor violent crimes, age 40 (-0.537**)	0.066 ... 0.371
# of felony arrests, age 40 (-0.383**)	0.050 ... 0.369
# of lifetime violent crimes, age 40 (-0.574**)	0.046 ... 0.320
Months in all marriages, age 40 (39.6*)	0.185 ... 0.224 ... 0.269 ... 0.352

Percent (0, 20, 40, 60, 80, 100)

Legend: ▨ Cognitive Factor ■ Externalizing Behavior ⧄ Academic Motivation ☐ Other Factors

Abecedarian

High School Grad at 30	0.50 ... 0.20
Treatment Effect: .189 (p-value: .027)	
Ever 4 Yr College at 30	0.52 ... 0.20 ... 0.28
Treatment Effect: .193 (p-value: .016)	
Employment at 30	0.12 ... 0.38 ... 0.56
Treatment Effect: .19 (p-value: .016)	

Axis: 0, .2, .4, .6, .8, 1

Legend:
☐ D Cognitive Level   ☐ if p-value < .10
▦ D Character Level   ▦ if p-value < .10
|||| D Unobserved Skills Level |||| if p-value < .10