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

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

Treatment → Earlier Skill → Later Skill → Adult Outcome
CCWD: Where I fit

Treatment ➔ Earlier Skill ➔ Later Skill ➔ Adult Outcome
CCWD: Where I fit

Treatment → Earlier Skill → Later Skill → Adult Outcome
CCWD: Where I fit

*TWARNING*

Context

Treatment → Earlier Skill → Later Skill → Adult Outcome

+ + + +
CCWD: Where I fit?

*WARNING*

Do

![Diagram]

- Treatment
- Earlier Skill
- Later Skill
- Adult Outcome
- Context

[Diagram showing the relationship between treatment, earlier skill, later skill, and adult outcome, with context influencing all stages.]
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

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

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)
Earlier Mathematics Achievement \[\rightarrow\] Controls: IQ, working memory, reading achievement, SES, others \[\rightarrow\] Later Mathematics Achievement

\[+\]
RESEARCH ARTICLE

Statistically Controlling for Confounding Constructs Is Harder than You Think

Jacob Westfall*, Tal Yarkoni
Earlier Mathematics Achievement

Later Mathematics Achievement

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

Other stuff

Boring, plausible alternative theory
Earlier Mathematics Achievement vs. Later Mathematics Achievement

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

Methods: can we do better than this?
Earlier Mathematics Achievement vs. Later Mathematics Achievement

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

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: \(0.35\)

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

Figure 5: Correlations inferred from MS path estimates in Table 1

Note: All 4th and 5th grade impacts are p > .05. All correlations and other impacts are p < .05. Impacts are rescaled to be 1.0 in the spring of pre-K. Right scale shows non-rescaled impacts. Vertical lines depict 95% confidence intervals.
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 Bailey, Nguyen, Jenkins, Domina, Clements, & Sarama (2016, Developmental Psychology)

From Smith, Cobb, Farran, Cordray, & Munter (2013, AERJ)
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)
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

From Bailey, Duncan, Odgers, & Yu (2017, Journal of Research on Educational Effectiveness)
The “right” kinds of skills?
The “right” kinds of skills?

• Can we identify skills that are:
  • Malleable through intervention,

From Bailey, Duncan, Odgers, & Yu (2017, Journal of Research on Educational Effectiveness)
The “right” kinds of skills?

• Can we identify skills that are:
  • Malleable through intervention,
  • Fundamental for success,

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

<table>
<thead>
<tr>
<th>Grade transition</th>
<th>Reading tests</th>
<th>Math tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Margin of error</td>
</tr>
<tr>
<td>Grade K–1</td>
<td>1.52</td>
<td>±0.21</td>
</tr>
<tr>
<td>Grade 1–2</td>
<td>0.97</td>
<td>±0.10</td>
</tr>
</tbody>
</table>

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)
## Average Annual Gain in Effect Size From Nationally Normed Tests

<table>
<thead>
<tr>
<th>Grade transition</th>
<th>Reading tests</th>
<th>Math tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Margin of error</td>
</tr>
<tr>
<td>Grade K–1</td>
<td>1.52</td>
<td>±0.21</td>
</tr>
<tr>
<td>Grade 1–2</td>
<td>0.97</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 2–3</td>
<td>0.60</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 3–4</td>
<td>0.36</td>
<td>±0.12</td>
</tr>
<tr>
<td>Grade 4–5</td>
<td>0.40</td>
<td>±0.06</td>
</tr>
</tbody>
</table>

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)
## Average Annual Gain in Effect Size From Nationally Normed Tests

<table>
<thead>
<tr>
<th>Grade transition</th>
<th>Reading tests</th>
<th>Math tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Margin of error</td>
</tr>
<tr>
<td>Grade K–1</td>
<td>1.52</td>
<td>±0.21</td>
</tr>
<tr>
<td>Grade 1–2</td>
<td>0.97</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 2–3</td>
<td>0.60</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 3–4</td>
<td>0.36</td>
<td>±0.12</td>
</tr>
<tr>
<td>Grade 4–5</td>
<td>0.40</td>
<td>±0.06</td>
</tr>
<tr>
<td>Grade 5–6</td>
<td>0.32</td>
<td>±0.11</td>
</tr>
<tr>
<td>Grade 6–7</td>
<td>0.23</td>
<td>±0.11</td>
</tr>
<tr>
<td>Grade 7–8</td>
<td>0.26</td>
<td>±0.03</td>
</tr>
</tbody>
</table>

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)
### Average Annual Gain in Effect Size From Nationally Normed Tests

<table>
<thead>
<tr>
<th>Grade transition</th>
<th>Reading tests</th>
<th>Math tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Margin of error</td>
</tr>
<tr>
<td>Grade K–1</td>
<td>1.52</td>
<td>±0.21</td>
</tr>
<tr>
<td>Grade 1–2</td>
<td>0.97</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 2–3</td>
<td>0.60</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 3–4</td>
<td>0.36</td>
<td>±0.12</td>
</tr>
<tr>
<td>Grade 4–5</td>
<td>0.40</td>
<td>±0.06</td>
</tr>
<tr>
<td>Grade 5–6</td>
<td>0.32</td>
<td>±0.11</td>
</tr>
<tr>
<td>Grade 6–7</td>
<td>0.23</td>
<td>±0.11</td>
</tr>
<tr>
<td>Grade 7–8</td>
<td>0.26</td>
<td>±0.03</td>
</tr>
<tr>
<td>Grade 8–9</td>
<td>0.24</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 9–10</td>
<td>0.19</td>
<td>±0.08</td>
</tr>
<tr>
<td>Grade 10–11</td>
<td>0.19</td>
<td>±0.17</td>
</tr>
<tr>
<td>Grade 11–12</td>
<td>0.06</td>
<td>±0.11</td>
</tr>
</tbody>
</table>

From Hill, Bloom, Black, & Lipsey (2008, *Child Development Perspectives*)
### Average Annual Gain in Effect Size From Nationally Normed Tests

*WARNING*

<table>
<thead>
<tr>
<th>Grade transition</th>
<th>Reading tests</th>
<th>Math tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Margin of error</td>
</tr>
<tr>
<td>Grade K–1</td>
<td>1.52</td>
<td>±0.21</td>
</tr>
<tr>
<td>Grade 1–2</td>
<td>0.97</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 2–3</td>
<td>0.60</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 3–4</td>
<td>0.36</td>
<td>±0.12</td>
</tr>
<tr>
<td>Grade 4–5</td>
<td>0.40</td>
<td>±0.06</td>
</tr>
<tr>
<td>Grade 5–6</td>
<td>0.32</td>
<td>±0.11</td>
</tr>
<tr>
<td>Grade 6–7</td>
<td>0.23</td>
<td>±0.11</td>
</tr>
<tr>
<td>Grade 7–8</td>
<td>0.26</td>
<td>±0.03</td>
</tr>
<tr>
<td>Grade 8–9</td>
<td>0.24</td>
<td>±0.10</td>
</tr>
<tr>
<td>Grade 9–10</td>
<td>0.19</td>
<td>±0.08</td>
</tr>
<tr>
<td>Grade 10–11</td>
<td>0.19</td>
<td>±0.17</td>
</tr>
<tr>
<td>Grade 11–12</td>
<td>0.06</td>
<td>±0.11</td>
</tr>
</tbody>
</table>

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?

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

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


- Adverse early context
  - Early harsh parenting
  - Poor school readiness
  - Conduct problems
  - School failure

- Adolescent violence
  - Deviant Peers
  - Low parent monitoring

Early life  -----------------  Adolescence
Possible *foot in the door* pathways

Can we avoid negative and promote positive developmental cascades?


- **Adverse early context**
  - Early harsh parenting
  - Poor school readiness
  - Conduct problems
  - School failure
- **Adolescent violence**
  - Deviant Peers
  - Low parent monitoring

**Problematic if these probabilities multiply.**
Adverse early context

Early harsh parenting

Poor school readiness

Conduct problems

School failure

Adolescent violence

Deviant Peers

Low parent monitoring

Can we avoid negative and promote positive developmental cascades?


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
# of misdemeanor violent crimes, age 27 (-0.423**)
# of felony arrests, age 27 (-0.269**)
Jobless for more than 1 year, age 27 (-0.292*)
Ever tried drugs other than alcohol or weed, age 27 (-0.227**)
# of misdemeanor violent crimes, age 40 (-0.537**)
# of felony arrests, age 40 (-0.383**)
# of lifetime violent crimes, age 40 (-0.574**)
Months in all marriages, age 40 (39.6*)

Cognitive Factor  Externalizing Behavior  Academic Motivation  Other Factors
Perry Preschool

Abecedarian
Building Blocks

Building Blocks + Follow Through

Treatment Effect (end of pre-K)
Treatment Effect (end of K)
Forgetting
Learning of Difficult Problems

Effect Size (Pre-K Control Group SD Units)

Group

Effect Size (Probability)